VECTOR ANALYSIS

AND THE

THEORY OF RELATIVITY

BY

FRANCIS D. MURNAGHAN, M.A. (N. U. I.), PH.D.
Associate Professor of Applied Mathematics, Johns Hopkins University.

BALTIMORE
THE JOHNS HOPKINS PRESS
1922

513.7
M97v

The Murray Printing Company
LETTERPRESS AND OFFSET
Wakefield Massachusetts

One of the most striking effects of the publication of Einstein's papers on generalized relativity and of the discussions which arose in connection with the subsequent astronomical observations was to make students of physics renew their study of mathematics. At first they attempted to learn simply the technique, but soon there was a demand to understand more; real mathematical insight was sought. Unfortunately there were no books available, not even papers.

Dr. Murnaghan's little book is a most successful attempt to supply what is a definite need. Every physicist can read it with profit. He will learn the meaning of a vector for the first time. He will learn methods which are available for every field of mathematical physics. He will see which of the processes used by Einstein and others are strictly mathematical and which are physical. Every chapter is illuminating, and the treatment of the subject is that of a student of mathematics and is not developed *ad hoc.* The extension of surface and line integrals is most interesting for physicists and the discussion of the space relations in a four-dimensional geometry is one most needed. This is specially true concerning the case of point-symmetry which forms the basis of Einstein's formulae for gravitation as applied to the solar system.

I feel personally that I owe to this book a great debt. I have read it with care and shall read it again. It has given me a definiteness of understanding which I never had before, and a vision of a field of knowledge which before was remote.

JOSEPH S. AMES.

JOHNS HOPKINS UNIVERSITY,
June 1, 1921.

CONTENTS

PAGE

CHAPTER THREE

The Metrical Concept

CHAPTER FOUR

The Resolution of Tensors

CHAPTER FIVE

Integral Invariants and Moving Circuits

CHAPTER SIX

The Absolute Differential Calculus

CHAPTER SEVEN

Problems in Relativity

PREFACE.

This monograph is the outcome of a short course of lectures delivered, during the summer of 1920, to members of the graduate department of mathematics of The Johns Hopkins University. Considerations of space have made it somewhat condensed in form, but it is hoped that the mode of presentation is sufficiently novel to avoid some of the difficulties of the subject. It is our opinion that it is to the physicist, rather than to the mathematician, that we must look for the conquest of the secrets of nature and so it is to the physicist that this little book is addressed. The progress in both subjects during the last half century has been so remarkable that we cannot hope for investigators like Kelvin and Helmholtz who are equally masters of either. But this makes it, all the more, the pleasure and duty of the mathematician to adapt his powerful methods to the needs of the physicist and especially to explain these methods in a manner intelligible to any one well grounded in Algebra and Calculus.

The rapid increase in the number of text books in mathematics has created a problem of selection. We have tried to confine our references to a few *good* treatises which should be accessible to every student of mathematics.

Ch. V should be omitted on a first reading. In fact it is quite independent of the rest of the book and will be of interest mainly to students of Hydrodynamics and Theoretical Electricity. There are several paragraphs in Ch. IV which may be passed over by those interested mainly in the application of the theory to the problems of relativity. For these we may be permitted to suggest, before taking up the subject matter of Chap. VII, a reference to an essay "The Quest of the Absolute"

which appeared in the Scientific American Monthly, March (1921), and was reprinted in the book "Relativity and Gravitation,"* Munn & Co. (1921). It may be useful to add the well-known advice of the French physicist, Arago—"When in difficulty, read on."

The manuscript of the book was sent to the printer in June, 1921, and its delay in publication has been due to difficulties in the printing business. In the meantime several important papers bearing on the Theory of Relativity have appeared; it will be sufficient to refer the reader to some significant notes by Painlevé in the Comptes Rendus of this year (1922). We are under a debt of gratitude to Dr. J. S. Ames for valuable advice and enlivening interest. And, in conclusion, we must thank the officials of The Johns Hopkins Press for their painstaking care in this rather difficult piece of printing.

F. D. M.

OMAGH, IRELAND.
June, 1922.

* Edited by J. Malcolm Bird.

VECTOR ANALYSIS AND THE THEORY OF RELATIVITY

INTRODUCTION

Vector Analysis owes its origin to the German mathematicians Mōbius* and Grassmann† and their contemporary Sir William Hamilton.‡ Since its introduction it has had a rather checkered career and it is only within comparatively recent times that it has become an integral part of any course in Theoretical Physics. It is well known that the subject was regarded with disfavor by many able physicists, among whom Sir William Thomson, afterwards Lord Kelvin, was probably the most prominent. The reason for this is, in our opinion, not hard to seek. Grassmann, who undoubtedly had a much clearer conception of the generality and power of his methods than most of his followers, expounded the subject in a very abstract manner in order not to lose this generality. Naturally enough his writings attracted little attention and when, some forty years later, Heaviside§ and others were earnestly trying to popularize the method they swung to the other extreme and, in attempting to give an intuitive definition of what a vector is, failed to convey a clear and comprehensive idea. Roughly speaking their definition was

* *Mōbius, A. F., Der barycentrische Calcul* (1827). Werke, Bd. 1, Leipzig (1885).

† *Grassmann, H., Ausdehnungslehre* (1844). Werke, Bd. 1, Leipzig (1894). Grassmann was particularly interested in the operations he could perform upon his "vectors" and not in the transformations of the components of these which occur when a change of "basis" or coordinate system is made. In this respect the point of view of his work will be found very different from that adopted here.

‡ *Hamilton, W., Elements of Quaternions.* Dublin Univ. Press (1899).

§ *Heaviside, O., Electromagnetic Theory,* Vol. 1, Ch. 3. London (1893).

that "a vector is a quantity which, in addition to the quality of having magnitude, has that of direction." The fault with this definition is, of course, that it fails to explain just what is meant by "having direction." That this idea requires explanation is clear when we realise that the simple operation of rotating a body around a definite line through a definite angle—which, à priori, "has direction" in the same sense that an angular velocity has—is not a vector whilst an angular velocity is. Then, again, endless trouble arises when vectors are introduced in a manner making it difficult to see their "direction" and even today some of the better text-books on the subject speak of "*symbolic* vectors" such as gradient, curl, etc., as if they are in any way different from other vectors. In 1901 Ricci and Levi-Civita* published an account of their investigations of "*The Absolute Differential Calculus*"—a kind of differentiation of vectors. This paper was written in a very condensed form and did not at once attract the notice of students of Theoretical Physics. It was only in 1916 when Einstein† called attention to the usefulness of the results in that paper that it received adequate recognition. However it seems to be the common opinion that the methods there dealt with (and often referred to as the "mathematics of relativity") are extremely difficult. It is the purpose of this account to lessen this difficulty by treating several points in a more elementary and natural manner. For example, in an interesting introduction to their paper, Ricci and Levi-Civita point out, as an instance of the power of their methods, that they can obtain easily, by means of their absolute differentiation, the transformation of Laplace's differential operator Δ_2—which in Cartesian coordinates takes the form

$$\Delta_2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

* *Ricci, G., and Levi-Civita, T. Méthodes de Calcul différentiel absolu.* Math. Annalen, Bd. 54, p. 125 (1901).

† *Einstein, A., Die Grundlage der allgemeinen Relativitätstheorie.* Annalen der Physik, Bd. 49, p. 169 (1916).

—into any curvilinear coordinates whatsoever. This transformation was first obtained by Jacobi,* and, while expressing admiration for the ingenuity of his method, they justly remark that it is not perfectly satisfactory for the reason that it brings in ideas—those of the Calculus of Variations—foreign to the nature of the problem. Now by a method due to Beltrami† it happens that this very transformation can be obtained by Vector Analysis without any knowledge of absolute differentiation; the apparently fortuitous and happy disappearance from the final result of the troublesome three index symbols of that part of the subject is thus explained. In addition we hope to make it clear that the methods of the "Mathematics of Relativity" are applicable to, and necessary for, Theoretical Physics in general and will abide even if the Theory of Relativity has to take its place with the rejected physical theories of the past.

* *Jacobi, C. G.*, Werke, Bd. 2, p. 191. Berlin (1882).

† *Beltrami*, Ricerche di analisi applicata alla geometria. *Giornale di matematiche* (1864), p. 365.

CHAPTER I

1. Every student of physics knows the important rôle played by line, surface and volume integrals in that subject. For example, the scalar magnitude *work* is the line integral of the scalar magnitude *force* and this will suggest a simple mode of defining a vector. As, however, we shall wish to apply our results in part to gravitational spaces it is desirable at the outset to state as clearly as possible what we mean by the various terms employed.

Space.—By this term is meant a continuous* arrangement or set of points; a point being merely a group of n ordered real numbers. In our applications n is either 1, 2, 3, or 4 and the space is said to be of one, two, three, or four dimensions respec- [illegible] ordered group of numbers we denote by $x^{(1)}$, $x^{(2)}$, [illegible], and call the coordinates of the point they define. Nothing need be said for the present as to what the coordinates actually signify. A space defined in this way is a very abstract mathematical idea and to distinguish it from a more concrete idea of space in which, in addition to the above, we have a fundamental concept called *length,* we may, where necessary, call the latter a metrical space and the former a non-metrical space. We use the symbol S_n to indicate our space, metrical or not, of n dimensions.

Spreads in S_n

It is possible to choose from the points of S_n an arrangement or set of points such that any one point is determined by the value of a single variable. Thus if, instead of being perfectly independent, the n coordinates $x^{(1)}, \cdots, x^{(n)}$ are all functions

*Continuity is assumed as an aid to mathematical treatment. In certain modern theories preference is given to a discontinuous or *discrete* set of points.

of a *single* independent variable, or parameter, u_1

$$x^{(s)} \equiv x^{(s)}(u_1) \qquad (s = 1, 2, \cdots, n)$$

the point x is said to trace a curve or spread of one dimension as u_1 varies continuously from the value u_1^0 to $u_1^{(1)}$. The points corresponding to the values $u_1 = u_1^0$ and $u_1 = u_1^{(1)}$ are called the end points of the curve and if they coincide, i.e., if all corresponding coordinates are equal the curve is said to be *closed*.

A spread of two dimensions in S_n is similarly defined by

$$x^{(s)} \equiv x^{(s)}(u_1, u_2) \qquad (s = 1, \cdots, n)$$

where u_1 and u_2 are independent parameters. Here we have two degrees of freedom because we can vary the point x by varying either u_1 or u_2. It is necessary, however, that the functions $x^s(u_1, u_2)$ should be distinct functions of the parameters u_1, u_2; the criterion for this being that not *all* the Jacobian determinants

$$\frac{\partial(x^{(s_1)}, x^{(s_2)})}{\partial(u_1, u_2)} \equiv \begin{vmatrix} \dfrac{\partial x^{(s_1)}}{\partial u_1} & \dfrac{\partial x^{(s_1)}}{\partial u_2} \\ \dfrac{\partial x^{(s_2)}}{\partial u_1} & \dfrac{\partial x^{(s_2)}}{\partial u_2} \end{vmatrix} \qquad \begin{pmatrix} s_1 = 1, \cdots, n \\ s_2 = 1, \cdots, n \end{pmatrix}$$

should vanish identically. If this were to happen, we would not have two degrees of freedom but only one and the points would lie on a curve and not on a proper spread of two dimensions.

Similarly by a spread of p dimensions in $S_n \cdots (p \leq n)$ we mean the locus of points x with p degrees of freedom;

$$x^{(s)} \equiv x^{(s)}(u_1, u_2, \cdots, u_p) \qquad (s = 1, \cdots, n)$$

where not all the Jacobian determinants

$$\frac{\partial(x^{(s_1)}, x^{(s_2)}, \cdots, x^{(s_p)})}{\partial(u_1, u_2, \cdots, u_p)} \qquad \begin{bmatrix} s_1 = 1, \cdots, n \\ s_2 = 1, \cdots, n \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ s_p = 1, \cdots, n \end{bmatrix}$$

vanish identically. This we denote by V_p (the corresponding French term being *variété*) and we shall suppose all our V_p to be "smooth"; by this we mean that all the partial derivatives

$$\frac{\partial x^{(s)}}{\partial u_m} \qquad \begin{pmatrix} s = 1, \cdots, n \\ m = 1, \cdots, p \end{pmatrix}$$

are continuous. This restriction is not really necessary but is made to avoid accessory difficulties.

Integral over a spread of one dimension V_1*

Consider an ordered set of n arbitrary continuous functions $X_1, \cdots, X_n$ of the coordinates $x^{(1)}, \cdots, x^{(n)}$. (For brevity sake we shall hereafter use the phrase "functions of position.") The numerical value assigned to the label r in the symbol X_r tells which one of the components $X_1, \cdots, X_n$, which are *ordered* or arranged in this sequence, we are discussing. Now for any curve V_1 given by

$$x^{(s)} \equiv x^{(s)}(u_1) \qquad (s = 1, \cdots, n)$$

n the differentials

$$dx^{(s)} \equiv \frac{\partial x^{(s)}}{\partial u_1} du_1 \qquad (s = 1, \cdots, n)$$

and then form the sum $X_1 dx^{(1)} + X_2 dx^{(2)} \cdots + X_n dx^{(n)}$ which is, by definition, identically the same as

$$\sum_{s=1}^{n} \left(X_s \frac{\partial x^{(s)}}{\partial u_1} \right) du_1$$

If in each of the functions X_s of position we replace the coordinates $x^{(1)}, \cdots, x^{(n)}$ by their values on the curve V_1

$$x^{(s)} \equiv x^{(s)}(u_1) \qquad (s = 1, \cdots, n)$$

$\sum_{s=1}^{n} X_s \frac{\partial x^{(s)}}{\partial u_1}$ becomes a function of u_1, $F(u_1)$ let us say, and we may

* Reference should be made to the classical paper by *H. Poincaré*, "Sur les résidus des intégrales doubles," Acta Math. (9), p. 321 (1887).

evaluate the definite integral $\int_{u_1^0}^{u_1^{(1)}} F(u_1)du_1$. This is called integral of the ordered set of n functions of position $(X_1, \cdots,$ over the curve. If, now, we change the parameter u_1 to some other parameter v_1 by means of the equation $u_1 = u_1(v_1)$ points on the curve are given by $x^{(s)} = x^{(s)}(u_1) = \bar{x}^{(s)}(v_1)$ $(s = 1, \cdots, n)$ and it is conceivable that the value of the integral might depend not only on the curve but on the parameter used in specifying the curve. However this is not the case since

$$\int_{u_1^0}^{u_1^{(1)}} F(u_1)du_1 \equiv \int_{u_1^0}^{u_1^{(1)}} \left\{ \sum_{s=1}^{n} X_s \frac{\partial x^{(s)}}{\partial u_1} \right\} du_1$$

and

$$\int_{v_1^0}^{v_1^{(1)}} \bar{F}(v_1)dv_1 \equiv \int_{v_1^0}^{v_1^{(1)}} \left\{ \sum_{s=1}^{n} X_s \frac{\partial \bar{x}^s}{\partial v_1} \right\} dv_1 \equiv \int_{v_1^0}^{v_1^{(1)}} \left\{ \sum_{s=1}^{n} X_s \frac{\partial x^{(s)}}{\partial u_1} \right\} \frac{du_1}{dv_1} dv_1$$
$$\equiv \int_{u_1^0}^{u_1^{(1)}} \left\{ \sum_{1}^{n} X_s \frac{\partial x^{(s)}}{\partial u_1} \right\} du_1$$

This independence, on the part of the integral, of the accidental parameter used in describing the curve allows us to speak of the integral as *attached* to the curve and the symbol $\int \sum_{s=1}^{n} X_s dx^{(s)}$ is used since it contains no reference to the parameter u.

In what follows we shall adopt the convention that when a literal label occurs twice in a term, summation with respect to that label over the values $1, \cdots, n$ is implied. Thus our line integral may be conveniently written

$$I_1 \equiv \int X_s dx^{(s)}$$

Such a label has been called by Eddington a dummy label (or symbol) of summation. We prefer to adopt the term " umbral " used by Sylvester in a similar connection; the word signifying that the symbol has merely a shadow-like significance disappearing, as it does, when the implied summation is performed.

2. Integral I_2 over a spread V_2 of two dimensions

Consider a set of n^2 ordered functions of position (to indicate which we use two labels s_1, s_2)

$$X_{s_1, s_2} \qquad (s_1, s_2 = 1, \cdots, n)$$

The numerical values assigned to s_1 and s_2 tell which one of the set of n^2 functions we wish to discuss. It is convenient to think of the functions as arranged in a square or "checkerboard" with n rows and n columns; then s_1 may indicate the row and s_2 the column. V_2 is specified by means of two parameters u_1, u_2 through the equations $x^{(s)} \equiv x^{(s)}(u_1, u_2)$. Substitute these expressions for the coordinates in the functions $X_{s_1s_2}$ and consider the definite double integral

$$I_2 = \int \left(X_{s_1s_2} \frac{\partial x^{(s_1)}}{\partial u_1} \frac{\partial x^{(s_2)}}{\partial u_2} \right) du_1 du_2 \qquad (s_1 \text{ and } s_2 \text{ umbral labels})$$

extended over the values of u_1, u_2 which specify the points of V_2. This integral will depend for its value not only on the spread V_2 but on the parameters u_1, u_2 used to specify it unless the set X_{s_1, s_2} is *alternating*, i.e., $X_{s_1, s_2} \equiv -X_{s_2, s_1}$ which implies the identical vanishing of the n functions $X_{1,1}; \cdots X_{n,n}$ and the in pairs of the remaining $n^2 - n$ so that $-1)/2$ distinct functions in the set. Grouping functions of each pair we have

$$I_2 = \int X_{s_1, s_2} \frac{\partial(x^{(s_1)}\ x^{(s_2)})}{\partial(u_1, u_2)} du_1 du_2 \qquad (s_1 < s_2)$$

where now the umbral symbols do not take independently all values from 1 to n but only those for which the numerical value of s_1 is less than that of s_2. If a change of parameters is made by means of the equations

$$u_1 \equiv u_1(v_1, v_2)$$
$$u_2 \equiv u_2(v_1, v_2)$$

where u_1 and u_2 are distinct functions of v_1 and v_2 the coordinates are given by equations

$$x^{(s)} \equiv x^{(s)}(u_1, u_2) \equiv \bar{x}^{(s)}(v_1, v_2) \qquad (s = 1, \cdots, n)$$

and the value of I_2 when the v_1, v_2 are used as parameters is

$$\int \left\{ X_{s_1s_2} \frac{\partial(\bar{x}^{(s_1)}\ \bar{x}^{(s_2)})}{\partial(v_1, v_2)} \right\} dv_1 dv_2 \qquad (s_1 < s_2)$$

which, by the rule for multiplying Jacobians,

$$\equiv \int \left\{ X_{s_1 s_2} \frac{\partial(x^{(s_1)}\, x^{(s_2)})}{\partial(u_1, u_2)} \right\} \frac{\partial(u_1, u_2)}{\partial(v_1, v_2)} dv_1 dv_2 \qquad (s_1 < s_2)$$

and this by the formula for the change of variables in a double integral

$$\equiv \int \left\{ X_{s_1 s_2} \frac{\partial(x^{(s_1)}, x^{(s_2)})}{\partial(u_1, u_2)} \right\} du_1 du_2 \qquad (s_1 < s_2)$$

Starting, then, with an *alternating* set of functions of position $X_{s_1 s_2}$ we can form an integral, (over any V_2), which depends in no way on the parameters chosen to specify it. To avoid all reference to the accidental parameters we write I_2 in the abbreviated form $\int \{X_{s_1, s_2} d(x^{(s_1)}, x^{(s_2)})\}$ $(s_1 < s_2)$. We adopt this in preference to the customary notation $\int \{X_{s_1 s_2} dx^{(s_1)} dx^{(s_2)}\}$ $(s_1 < s_2)$ since no *product of differentials*, such as will occur later when we use quadratic differential forms, is implied.

In an exactly similar way an integral I_p over a spread V_p of p dimensions $(p \leq n)$ is defined.* By an alternating set of functions $X_{s_1, s_2, \cdots, s_p}$ of position we mean that a single interchange of two of the labels merely changes the sign of the function. If, then, two of these labels are the same the function must be identically zero. Then

$$I_p \equiv \int \left\{ X_{s_1, s_2, \cdots, s_p} \frac{\partial x^{(s_1)}}{\partial u_1} \cdots \frac{\partial x^{(s_p)}}{\partial u_p} \right\} du_1 \cdot du_2 \cdots du_p$$

is a definite multiple integral of order p extended over the values of $u_1, \cdots, u_p$ which specify the points of V_p. We write

$$I_p \equiv \int \left\{ X_{s_1, \cdots, s_p} \frac{\partial(x^{s_1}, \cdots, \partial x^{(s_p)})}{\partial(u_1, \cdots, u_p)} \right\} du_1 \cdots du_p \quad (s_1 < s_2 \cdots s_p)$$

where, in the summation with respect to the umbral symbols, $s_1, s_2, \cdots, s_p$, $s_1 < s_2 < \cdots < s_p$. To emphasize the fact that

* When $p = n$ it is customary to use the phrase *region of* S_n in preference to spread of n dimensions in S_n.

I_p does not depend in any way on the parameters $u_1, \cdots, u_p$ it will be written

$$I_p \equiv \int X_{s_1, \cdots, s_p} d(x^{(s_1)}, \cdots, x^{(s_p)}) \qquad (s_1 < s_2 \cdots < s_p)$$

Examples. $n = 4 \quad x^1 = x, \; x^{(2)} = y, \; x^{(3)} = z, \; x^{(4)} = t$

$$X_1 = X, X_2 = Y, \text{etc.}$$

$$I_1 \equiv \int (Xdx + Ydy + Zdz + Tdt)$$

$$I_2 \equiv \int X_{23} d(y, z) + X_{3,1} d(z, x) + X_{12} d(x, y) + X_{14} d(x, t) + X_{24} d(y, t) + X_{34} d(z, t)$$

$$I_3 \equiv \int X_{123} d(x, y, z) + X_{124} d(x, y, t) + X_{134} d(x, z, t) + X_{234} d(y, z, t)$$

$$I_4 \equiv \int X_{1, 2, 3, 4} d(x, y, z, t)$$

Here in I_2 we may write $X_{31} d(z, x)$ instead of $X_{1,3} d(x, z)$ since

$$X_{31} = - X_{13} \quad \text{and} \quad d(z, x) = - d(x, z)$$

example of I_2 we may take the case of a moving y Euclidean space of three dimensions, the curve ... to change in a continuous manner as it moves. ...ere x, y, z may be rectangular Cartesian coordinates and t may denote the Newtonian time. u_1 is any parameter which serves to locate the points of the curve at any definite time $t = t_0$ and u_2 may well be taken $\equiv t$. Then the equations of our V_2 are

$$x \equiv x(u_1, t); \quad y \equiv y(u_1, t); \quad z \equiv z(u_1, t); \quad t \equiv u_2$$

and the *parameter curves* $u_2 =$ constant are the various positions of the moving curve, whilst the curves $u_1 =$ constant are the paths of definite points on the initial position of the moving curve. Denote $\partial x/\partial t$ by $\dot{x}$ and we have

$$d(y, z) \equiv \frac{\partial(y, z)}{\partial(u_1, t)} du_1 dt \equiv \left(\frac{\partial y}{\partial u_1} \dot{z} - \frac{\partial z}{\partial u_1} \dot{y} \right) du_1 dt$$

$$d(x, t) \equiv \frac{\partial(x, t)}{\partial(u_1, t)} du_1 dt \equiv \frac{\partial x}{\partial u_1} \cdot du_1 dt$$

(It may not be superfluous to point out that it is essential to the argument that u_1 and u_2 should be *independent* variables. Thus in the present example u_1 could not stand for the arc distance from an end point of the moving curve *if the curve deforms as it moves* although it could conveniently stand for the initial arc distance.) Our I_2 may here be written

$$\int \Big\{ (X_{14} + X_{12}\dot{y} - X_{31}\dot{z}) \frac{\partial x}{\partial u_1} + (X_{24} + X_{23}\dot{z} - X_{12}\dot{x}) \frac{\partial y}{\partial u_1} + (X_{34} + X_{31}\dot{x} - X_{23}\dot{y}) \frac{\partial z}{\partial u_1} \Big\} du_1 dt$$

showing it in the form of a time integral of a certain line integral taken over the moving curve. Before proceeding to define the idea of vector quantities it is necessary to make one remark of a physical nature. We have written expressions of the type

$$X_s dx^{(s)} \qquad (s \text{ an umbral symbol})$$

and regarded the separate terms of these expressions $X_1 dx^{(1)}, \cdots,$ etc., as mere numbers. To actually perform the indicated summations it is necessary, when we apply our methods to physics, that the separate terms in a summation should be of the same kind, i.e., *have the same dimensions.* Thus if the coordinates $x^{(1)} \cdots x^{(n)}$ are all of the same kind the coefficients

$$X_{s_1, \ldots, s_p} \qquad \left\{ \begin{matrix} s_1 \\ \vdots \\ s_p \end{matrix} = 1, \cdots, n \right\}$$

occurring in the various integrals must all have the same dimensions.

3. Transformation of Coordinates

It has already been seen that if the various line integrals under discussion are to have values independent of the choice of *parameters* $(u_1, \cdots, u_p)$ care must be taken that the n^p functions of position $X_{s_1, \ldots, s_p}$ which form the coefficients of the I_p should

be alternating. Let us now see what happens to these coefficients when we change, for some reason, the coordinates $x^{(1)}, \cdots, x^{(n)}$ used to specify the points of the V_p. The formulæ of transformation are given by n equations

$$x^{(s)} \equiv x^{(s)}(y^{(1)}, \cdots, y^{(n)}) \qquad (s = 1, \cdots, n)$$

the functions $x^{(s)}$ being supposed distinct so that the Jacobian of the transformation

$$J \equiv \frac{\partial(x^{(1)}, \cdots, x^{(n)})}{\partial(y^{(1)}, \cdots, y^{(n)})}$$

does not vanish identically. These equations may be regarded in two ways. First the $y^{(s)}$ may each denote the same idea as the corresponding $x^{(s)}$ and then we have a correspondence set up between a point y and some, in general different, point x. Secondly the symbols $y^{(s)}$ may have a meaning quite distinct from the symbols $x^{(s)}$ and then we have a correspondence between one set of coordinates $y^{(s)}$ of a point and another set coordinates $x^{(s)}$ of the *same point*. It is the second way of ...ing at the matter that interests us and we speak then of a transformation of *coordinates*. (From the first point of view we would have a *point correspondence*.) Since the functions $x^{(s)}$ are distinct we can, in general, solve the equations* and obtain

$$y^{(s)} \equiv y^{(s)}(x^{(1)}, \cdots, x^{(n)}) \qquad (s = 1, \cdots, n)$$

As an example take $n = 3$ and let $x^{(1)}, x^{(2)}, x^{(3)}$ be rectangular Cartesian coordinates and $(y^{(1)}, y^{(2)}, y^{(3)})$ space polar coordinates in ordinary Euclidean space of three dimensions.

$x^{(1)} \equiv y^{(1)} \sin y^{(2)} \cos y^{(3)}$	$y^{(1)} \equiv +\sqrt{x^{(1)2} + x^{(2)2} + x^{(3)2}}$
$x^{(2)} \equiv y^{(1)} \sin y^{(2)} \sin y^{(3)}$	$y^{(2)} \equiv \tan^{-1}\sqrt{\dfrac{x^{(1)2} + x^{(2)2}}{x^{(3)}}}$
$x^{(3)} \equiv y^{(1)} \cos y^{(2)}$	$y^{(3)} \equiv \tan^{-1}\dfrac{x^{(2)}}{x^{(1)}}$

* Cf. *Goursat-Hedrick*, Mathematical Analysis, Vol. 1, Ch. 2, or *Wilson, E. B.*, Advanced Calculus.

In order to have a *uniform* transformation of coordinates—so that to a given set of numbers $y^{(1)}$, $y^{(2)}$, $y^{(3)}$ there may correspond but one set $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ and conversely—it is frequently necessary to restrict the range of values of one or the other set. Thus in the example chosen we put $y^{(1)} \geq 0$; $0 \leq y^{(2)} \leq \pi$; $0 \leq y^{(3)} < 2\pi$. If now in

$$I_1 \equiv \int X_s dx^{(s)} \qquad (s \text{ an umbral symbol})$$

we substitute

$$x^{(s)} \equiv x^{(s)}(y^{(1)}, \cdots, y^{(n)}) \qquad (s = 1, \cdots, n)$$

X_s becomes $\overline{X}_s(y^1, \cdots, y^n)$ say, and $dx^{(s)} \equiv \dfrac{\partial x^{(s)}}{\partial u_1} du_1$ becomes

$$\left(\frac{\partial x^{(s)}}{\partial y^{(r)}} \frac{\partial y^{(r)}}{\partial u_1}\right) du_1 \qquad (r \text{ an umbral symbol})$$

and so I_1 becomes

$$\int \left(\overline{X}_s \frac{\partial x^{(s)}}{\partial y^{(r)}} \frac{\partial y^{(r)}}{\partial u_1}\right) du_1 \qquad (r, s \text{ both umbral symbols})$$

$$\equiv \int \overline{X}_s \frac{\partial x^{(s)}}{\partial y^{(r)}} dy^{(r)} \equiv \int Y_r dy^{(r)}$$

where Y is defined by the equation

$$Y_r \equiv \overline{X}_s \frac{\partial x^{(s)}}{\partial y^{(r)}} \qquad (r = 1, \cdots, n;\ s \text{ umbral})$$

We shall from this on drop the bar notation above the X_s which indicates that the substitution $x^{(s)} \equiv x^{(s)}(y^{(1)}, \cdots, y^{(n)})$ has been carried out. It will always be clear when this is supposed done. For an I_2 we have

$$I_2 \equiv \int X_{s_1 s_2} d(x^{(s_1)}, x^{(s_2)}) \qquad (s_1 < s_2) \qquad (s_1, s_2 \text{ umbral})$$

$$\equiv \int \left\{X_{s_1 s_2} \frac{\partial(x^{s_1}, x^{s_2})}{\partial(u_1, u_2)}\right\} du_1 du_2 \qquad (s_1 < s_2) \text{ by definition}$$

$$\equiv \int \left\{X_{s_1 s_2} \frac{\partial x^{(s_1)}}{\partial u_1} \frac{\partial x^{(s_2)}}{\partial u_2}\right\} du_1 du_2$$

since the functions $X_{s_1 s_2}$ form an alternating set.

Now

$$\frac{\partial x^{(s_1)}}{\partial u_1} \equiv \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \frac{\partial y^{(r_1)}}{\partial u_1} \qquad (r_1 \text{ umbral})$$

so that

$$\frac{\partial x^{(s_1)}}{\partial u_1} \cdot \frac{\partial x^{(s_2)}}{\partial u_2} \equiv \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \frac{\partial x^{(s_2)}}{\partial y^{(r_2)}} \frac{\partial y^{(r_1)}}{\partial u_1} \frac{\partial y^{(r_2)}}{\partial u_2}$$

(r_1 and r_2 both umbral symbols)

Hence if we define

$$Y_{r_1 r_2} \equiv X_{s_1 s_2} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \frac{\partial x^{(s_2)}}{\partial y^{(r_2)}} \qquad (s_1 \text{ and } s_2 \text{ umbral})$$

I_2 takes the form

$$\int \left\{ Y_{r_1 r_2} \frac{\partial y^{(r_1)}}{\partial u_1} \frac{\partial y^{(r_2)}}{\partial u_2} \right\} du_1 du_2$$

Now

$$\frac{\partial x^{(s_1)}}{\partial y^{(r_2)}} \frac{\partial x^{(s_2)}}{\partial y^{(r_1)}} \text{ (by definition)}$$

$$\equiv X_{s_2, s_1} \frac{\partial x^{(s_2)}}{\partial y^{(r_2)}} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}}$$ (by a mere interchange of the letters standing for the umbral symbols s_1 and s_2)

$$\equiv - X_{s_1 s_2} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \frac{\partial x^{(s_2)}}{\partial y^{(r_2)}}$$ (since X_{s_1, s_2} is alternating by definition)

$$\equiv - Y_{r_1, r_2} \text{ (by definition)}$$

Accordingly the set of functions Y_{r_1, r_2} of position, defined as above, is also alternating and we may write

$$I_2 \equiv \int Y_{r_1 r_2} d(y^{(r_1)}, y^{(r_2)}) \qquad (r_1 < r_2)$$

Generalizing we may write I_p in the form

$$\int Y_{r_1, \ldots, r_p} d(y^{(r_1)}, \cdots, y^{(r_p)})$$

($r_1, \cdots, r_p$ umbral) and $r_1 < r_2 < \cdots < r_p$

where the coefficients $Y_{r_1, \ldots, r_p}$ form an alternating set of n^p functions of position defined by the equations

$$Y_{r_1, \ldots, r_p} \equiv X_{s_1, \ldots, s_p} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \cdots \frac{\partial x^{(s_p)}}{\partial y^{(r_p)}} \quad (s_1, \ldots, s_p \text{ umbral symbols})$$

Accordingly, then, if an integral over a curve, or more generally a spread of dimensions p, is to have a value independent of the coordinates the coefficients are completely determined in *every system of coordinates* once they are known in any particular system of coordinates. The coefficients in a line integral form as we shall see later a set of functions which "have direction" in Heaviside's sense and so might be called a vector. As, however, the term vector is derived from a geometrical interpretation of the idea which loses to a great extent its significance when we apply our ideas to spaces of arbitrary metrical character the name has been changed and the coefficients of a line integral are said to form, taken as a group, a Tensor of the first rank of which the coefficients are the ordered *components*.* To distinguish between this definition and another of similar character this Tensor is said to be *covariant*. More generally the coefficients of an I_p, n^p in number, are said to form a covariant tensor of rank p of which the separate coefficients $X_{s_1, \ldots, s_p}$ are the ordered components. Knowing the values of the components $X_{s_1, \ldots, s_p}$ of a covariant tensor in any suitable system of coordinates $x^{(s)}$ the components in any other set $y^{(s)}$ are furnished by the equations

$$Y_{r_1, \ldots, r_p} \equiv X_{s_1, \ldots, s_p} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \cdots \frac{\partial x^{(s_p)}}{\partial y^{(r_p)}} \quad (s_1, \cdots, s_p \text{ umbral labels})$$

Although not of such physical importance it is convenient to extend the idea of Tensor to an *arbitrary* set of functions of position $X_{s_1, \ldots, s_p}$ which follow the same law of correspondence, when a transformation of coordinates is made, as the *alternating* set above. If we do this it is merely the *alternating* covariant Tensors which arise as coefficients in integrals over geometric figures. The *reason* for the correspondence between the com-

* The term Tensor was used by Gibbs in another sense in his lectures (see his *Vector Analysis*, Chap. V, edited by Wilson, E. B.) and also with the same meaning as that given here by Voigt, W., "Die fundamentalen Eigenschaften der Krystalle," Leipzig (1898). Cf. Ch. IV, § 4, infra.

ponents in different systems of a Tensor in the general non-alternating case would remain to be explained.

4. Introduction of contravariant tensors

In the expression

$$I_1 \equiv \int X_s dx^{(s)} \equiv \int Y_s dy^{(s)} \qquad (s \text{ umbral})$$

the quantities by which the components X_s of the covariant tensor of rank one are multiplied have a law of correspondence defined by the equations

$$dy^{(s)} \equiv \frac{\partial y^{(s)}}{\partial u} \cdot du \equiv \left(\frac{\partial y^{(s)}}{\partial x^{(r)}} \cdot \frac{\partial x^{(r)}}{\partial u}\right) du \qquad (r \text{ umbral})$$

$$\equiv \frac{\partial y^{(s)}}{\partial x^{(r)}} \cdot dx^{(r)}$$

Similarly in the integral

$$I_2 \equiv \int \left(X_{rs} \frac{\partial x^{(r)}}{\partial u_1} \frac{\partial x^{(s)}}{\partial u_2}\right) du_1 du_2 \equiv \int \left(Y_{rs} \frac{\partial y^{(r)}}{\partial u_1} \frac{\partial y^{(s)}}{\partial u_2}\right) du_1 du_2$$

the factors X^{rs}, Y^{rs} which multiply the components X_{rs}, Y_{rs} respectively of the alternating covariant tensor of rank two have a law of correspondence given by the equations

$$Y^{s_1 s_2} \equiv \frac{\partial y^{(s_1)}}{\partial u_1} \frac{\partial y^{(s_2)}}{\partial u_2} du_1 du_2 \qquad \text{(by definition)}$$

$$\equiv \frac{\partial y^{(s_1)}}{\partial x^{(r_1)}} \frac{\partial x^{(r_1)}}{\partial u_1} \cdot \frac{\partial y^{(s_2)}}{\partial x^{(r_2)}} \frac{\partial x^{(r_2)}}{\partial u_2} \cdot du_1 du_2 \qquad (r_1, r_2 \text{ umbral symbols})$$

$$\equiv \frac{\partial y^{(s_1)}}{\partial x^{(r_1)}} \cdot \frac{\partial y^{(s_2)}}{\partial x^{(r_2)}} \cdot X^{r_1 r_2} \qquad \text{(by definition)}$$

and so in general for an integral over a spread of p dimensions ($p \leq n$). These factors, regarded as a whole, are said to form a *contravariant Tensor* of the first, second, $\cdots$, pth rank as the case may be. The sets introduced in this way are *not*, as in the case of the covariant tensors, alternating. Even though the correspondence between the two sets of functions of position

$X^{s_1 s_2 \cdots s_p}$ and $Y^{s_1 \cdots s_p}$ may not *arise* in the above manner the set is said to form a contravariant tensor of rank p if the correspondence between the *ordered* components is defined by the equations

$$Y^{s_1, \cdots, s_p} \equiv X^{r_1, \cdots, r_p} \frac{\partial y^{(s_1)}}{\partial x^{(r_1)}} \cdots \frac{\partial y^{(s_p)}}{\partial x^{(r_p)}} \quad (r_1, \cdots, r_p \text{ umbral})$$

The labels which serve to order the components are written *above* in the case of contravariant and *below* in the case of covariant Tensors. The following remark may be useful in aiding the beginner to remember easily the important equations defining the correspondence. The umbral symbols are always attached to the x coordinates on the right. When the labels are $\left.\begin{matrix}\text{below}\\ \text{above}\end{matrix}\right\}$ on the left the y coordinates are $\left.\begin{matrix}\text{below}\\ \text{above}\end{matrix}\right\}$ on the right.

Thus

$$Y_{r_1 r_2} \equiv X_{s_1 s_2} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \frac{\partial x^{(s_2)}}{\partial y^{(r_2)}} \qquad (s_1, s_2 \text{ umbral})$$

whilst

$$Y^{r_1 r_2} \equiv X^{s_1 s_2} \frac{\partial y^{(r_1)}}{\partial x^{(s_1)}} \frac{\partial y^{(r_2)}}{\partial x^{(s_2)}} \qquad (s_1, s_2 \text{ umbral})$$

By an obvious and useful extension we can now introduce *mixed Tensors* partly covariant and partly contravariant in nature. Thus the set of n^3 functions of position $X_{r_1 r_2}^{r_3}$ form a mixed tensor of rank three, covariant of rank two and contravariant of rank one, if the correspondence between the two sets of ordered components is defined by the equations

$$Y_{r_1 r_2}^{r_3} \equiv X_{s_1 s_2}^{s_3} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \frac{\partial x^{(s_2)}}{\partial y^{(r_2)}} \frac{\partial y^{(r_3)}}{\partial x^{(s_3)}} \quad (s_1, s_2, s_3 \text{ umbral symbols})$$

Now when we recall that the x coordinates are perfectly arbitrary as also are the y's it becomes apparent that it must be possible to interchange the x and y coordinates in the equations

defining the correspondence. Thus, to give a concrete example, it must be possible to derive from the n^2 equations

$$Y_{r_1, r_2} \equiv X_{s_1 s_2} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \frac{\partial x^{(s_2)}}{\partial y^{(r_2)}} \qquad (s_1, s_2 \text{ umbral})$$

which serve to define a covariant tensor of rank 2, the equations

$$X_{r_1 r_2} \equiv Y_{s_1 s_2} \frac{\partial y^{(s_1)}}{\partial x^{(r_1)}} \frac{\partial y^{(s_2)}}{\partial x^{(r_2)}} \qquad (s_1, s_2 \text{ umbral})$$

In fact

$$Y_{s_1 s_2} \frac{\partial y^{(s_1)}}{\partial x^{(r_1)}} \frac{\partial y^{(s_2)}}{\partial x^{(r_2)}} \equiv X_{t_1 t_2} \frac{\partial x^{(t_1)}}{\partial y^{(s_1)}} \frac{\partial x^{(t_2)}}{\partial y^{(s_2)}} \frac{\partial y^{(s_1)}}{\partial x^{(r_1)}} \frac{\partial y^{(s_2)}}{\partial x^{(r_2)}}$$

$$(s_1, s_2, t_1, t_2 \text{ all umbral})$$

$$\equiv X_{r_1 r_2}$$

$\Bigg[$for $\dfrac{\partial x^{(t_1)}}{\partial y^{(s_1)}} \dfrac{\partial y^{(s_1)}}{\partial x^{(r_1)}}$ (where s_1 is umbral) is $\equiv \dfrac{\partial x^{(t_1)}}{\partial x^{(r_1)}}$ by the rule for composite differentiation and this, on account of the mutual independence of the x coordinates, is $\equiv 0$ unless $t_1 = r_1$ in which case it $= 1\Bigg]$.

To conclude these definitions it will be sufficient to state that a single function of position may be regarded as a tensor *of rank zero* if its value (not its formal expression) is the same in all sets of coordinates. No labels are here required to order the components and the equation defining the correspondence is simply

$$Y \equiv X$$

Such a function of position is also called an *invariant* or *absolute* (or in the text-books on vector analysis a *scalar*) quantity. The reason for regarding this as a tensor (of either kind) of rank zero will become apparent from a study of the rules of operation with tensors.

Example.—Consider the formulæ of transformation from rectangular Cartesian to space polar coordinates (§ 3).

Here

$$\frac{\partial x^{(1)}}{\partial y^{(1)}} = \sin y^{(2)} \cos y^{(3)}; \qquad \frac{\partial x^{(1)}}{\partial y^{(2)}} = + y^{(1)} \cos y^{(2)} \cos y^{(3)};$$

$$\frac{\partial x^{(1)}}{\partial y^{(3)}} = - y^{(1)} \sin y^{(2)} \sin y^{(3)},$$

etc., and we obtain

$$\begin{aligned}
Y_1 &\equiv X_1 \frac{\partial x^{(1)}}{\partial y^{(1)}} + X_2 \frac{\partial x^{(2)}}{\partial y^{(1)}} + X_3 \frac{\partial x^{(3)}}{\partial y^{(1)}} \\
&\equiv (X_1 \sin y^{(2)} \cos y^{(3)} + X_2 \sin y^{(2)} \sin y^{(3)} + X_3 \cos y^{(2)}) \\
Y_2 &\equiv X_1 \frac{\partial x^{(1)}}{\partial y^{(2)}} + X_2 \frac{\partial x^{(2)}}{\partial y^{(2)}} + X_3 \frac{\partial x^{(3)}}{\partial y^{(2)}} \\
&\equiv y^{(1)}[X_1 \cos y^{(2)} \cos y^{(3)} + X_2 \cos y^{(2)} \sin y^{(3)} - X_3 \sin y^{(2)}] \\
Y_3 &\equiv X_1 \frac{\partial x^{(1)}}{\partial y^{(3)}} + X_2 \frac{\partial x^{(2)}}{\partial y^{(3)}} + X_3 \frac{\partial x^{(3)}}{\partial y^{(3)}} \\
&\equiv y^{(1)}[- X_1 \sin y^{(2)} \sin y^{(3)} + X_2 \sin y^{(2)} \cos y^{(3)}]
\end{aligned}$$

the X's on the right hand side being supposed expressed in terms of the y's. If then we denote by R, Θ, Φ the resolved parts of the vector X_1, X_2, X_3 (the theory of the resolution of tensors will be dealt with later but we may anticipate here) along the three polar coordinate directions at any point

$$Y_1 \equiv R; \qquad Y_2 \equiv y^{(1)}\Theta \equiv r\Theta; \qquad Y_3 \equiv y^{(1)} \sin y^{(2)}\Phi \equiv r \sin \theta\Phi$$

For a contravariant tensor of rank one we have

$$\begin{aligned}
Y^{(1)} &\equiv X^{(1)} \frac{\partial y^{(1)}}{\partial x^{(1)}} + X^{(2)} \frac{\partial y^{(1)}}{\partial x^{(2)}} + X^{(3)} \frac{\partial y^{(1)}}{\partial x^{(3)}} \\
&\equiv (X^{(1)} \sin y^{(2)} \cos y^{(3)} + X^{(2)} \sin y^{(2)} \sin y^{(3)} + X^{(3)} \cos y^{(2)}) \\
Y^{(2)} &\equiv X^{(1)} \frac{\partial y^{(2)}}{\partial x^{(1)}} + X^{(2)} \frac{\partial y^{(2)}}{\partial x^{(2)}} + X^{(3)} \frac{\partial y^{(2)}}{\partial x^{(3)}} \\
&= \frac{1}{y^{(1)}} (X^{(1)} \cos y^{(2)} \cos y^{(3)} + X^{(2)} \cos y^{(2)} \sin y^{(3)} - X^{(3)} \sin y^{(2)}) \\
Y^{(3)} &\equiv X^{(1)} \frac{\partial y^{(3)}}{\partial x^{(1)}} + X^{(2)} \frac{\partial y^{(3)}}{\partial x^{(2)}} + X^{(3)} \frac{\partial y^{(3)}}{\partial x^{(3)}}
\end{aligned}$$

$$\equiv \frac{1}{y^{(1)} \sin y^{(2)}} (- X^1 \sin y^{(3)} + X^2 \cos y^{(3)})$$

where the X's on the right are supposed expressed in terms of the y's. Call the resolved parts of $(X^{(1)}, X^{(2)}, X^{(3)})$ along the polar coordinate directions R, Θ, Φ as before and we have

$$Y^{(1)} \equiv R; \qquad Y^{(2)} \equiv \frac{\Theta}{r}; \qquad Y^{(3)} \equiv \frac{\Phi}{r \sin \theta} \; *$$

* A general result of which this is a special case is given in Chapter IV.

CHAPTER II

THE ALGEBRA OF TENSORS

1. Elementary rules for deriving and operating with tensors

(a) *The Rule of Linear Combination*

If $X^{r_1 \cdots r_q}_{s_1 \cdots s_p}$ is a tensor of rank $p + q \begin{pmatrix} p = 0, 1, \cdots \\ q = 0, 1, \cdots \end{pmatrix}$ and $\overline{X}^{r_1 \cdots r_q}_{s_1 \cdots s_p}$ is another tensor of the same kind then the set of n^{p+q} functions of position found by adding components of like order (that is with all corresponding labels, both upper and lower, having the same numerical values each to each) forms a tensor of the same kind as X and $\overline{X}$ which is called the sum of X and $\overline{X}$. By the phrase "of the same kind" we imply not only that X and $\overline{X}$ must have the same rank both as to covariant and contravariant character, but that corresponding components have the same dimensions. The proof of the statement is immediate for from the equations

$$Y^{\rho_1 \cdots \rho_q}_{\sigma_1 \cdots \sigma_p} = X^{r_1 \cdots r_q}_{s_1 \cdots s_p} \frac{\partial y^{(\rho_1)}}{\partial x^{(r_1)}} \cdots \frac{\partial y^{(\rho_q)}}{\partial x^{(r_q)}} \frac{\partial x^{(s_1)}}{\partial y^{(\sigma_1)}} \cdots \frac{\partial x^{(s_p)}}{\partial y^{(\sigma_p)}}$$

$\left({}^{r_1 \cdots r_q}_{s_1 \cdots s_p}\right.$ all umbral)

and a similar one obtained by writing a bar over Y and X we obtain by addition

$$(Y^{\rho_1 \cdots \rho_q}_{\sigma_1 \cdots \sigma_p} + \overline{Y}^{\rho_1 \cdots \rho_q}_{\sigma_1 \cdots \sigma_p}) = (X^{r_1 \cdots r_q}_{s_1 \cdots s_p} + \overline{X}^{r_1 \cdots r_q}_{s_1 \cdots s_p}) \frac{\partial y^{(\rho_1)}}{\partial x^{(r_1)}} \cdots \frac{\partial x^{(s_p)}}{\partial y^{(\sigma_p)}}$$

which is the mathematical formulation of the statement that $X + \overline{X}$ is a tensor of the same kind as both X and $\overline{X}$.

If we multiply the equations written above, which express the tensor character of $X^{r_1 \cdots r_q}_{s_1 \cdots s_p}$ by an invariant function of position

(possibly a constant) m we have that mX is a tensor of the same character as X. Combining this with the previous definition of a sum, repeatedly applied if necessary, we have what is known as a linear combination of Tensors

$$l_1X + l_2X^1 + \cdots$$

where the l_1, l_2, $\cdots$ are either mere numbers or scalar (invariant) functions. *The separate members of this linear combination must be of the same kind.* If, as a special case, l_2 is a negative number $l_2 = -1$ say and $l_1 = +1$ then $X + (-X^1)$ is written $X - X$ and in this way subtraction is defined. A tensor *all* of whose components are zero is said to be the zero tensor. (It is important to notice that the property of having all the components zero is an *absolute* one; i.e., it is independent of the particular choice of coordinates in terms of which the components are expressed. This follows at once from the equations defining the correspondence between the ordered components in different systems of coordinates. The General Principle of Relativity merely says that all physical laws may be expressed each by the vanishing of a certain tensor. This satisfies the necessary demand that the content of a physical law must be independent of the coordinates used to express it mathematically. The fixing of the number of dimensions n as 4 rather than 3 and the interpretation of the physical significance of the coordinates are the difficult parts of the theory of relativity; the demand that all physical laws express the equality of tensors has nothing to do with these and must be granted by everyone. Here we regard an invariant as a tensor of zero rank.) Since the idea of a linear combination of tensors is reducible to a linear combination of the corresponding components it follows that the order of the separate members in a linear combination is unimportant.

2. (*b*) *The Rule of Interchange of Order of Components.*

A specific example will show most briefly and clearly what is meant by this rule. Consider the covariant tensor $X_{r_1r_2}$ of the

second rank. The components have a definite order which may be conveniently specified by a square arrangement.

X_{11}	$X_{12} \cdots$	X_{1n}
X_{21}		
X_{n1}		X_{nn}

If now we rearrange the n^2 functions amongst the n^2 small squares in such a way that the rows and columns are interchanged, then this same interchange of rows and columns will take place in the square for any other coordinate system y. We denote the new ordered set by a bar thus—

$$\overline{X}_{r,s} \equiv X_{s,r} \qquad (r, s = 1, 2, \cdots, n)$$

From $\overline{X}_{r,s}$ we obtain $\overline{Y}_{rs}$ by means of the equations of correspondence and we wish to show that $\overline{Y}_{rs} \equiv Y_{sr}$ where the Y_{rs} are obtained from the X_{rs} by the same equations of correspondence.

All we have done is to rearrange the order of summation on the right hand side of the equations of correspondence and the formal proof is very easy.

$$\overline{Y}_{rs} \equiv \overline{X}_{\rho\sigma} \frac{\partial x^{(\rho)}}{\partial y^{(r)}} \frac{\partial x^{(\sigma)}}{\partial y^{(s)}} \qquad \text{by definition } (\rho \text{ and } \sigma \text{ umbral})$$

$$\equiv X_{\sigma\rho} \frac{\partial x^{(\rho)}}{\partial y^{(r)}} \frac{\partial x^{(\sigma)}}{\partial y^{(s)}} \qquad \text{from definition of } \overline{X}$$

$$\equiv Y_{sr} \qquad \text{(from equations of correspondence).}$$

Combining this rule with rule (a) we derive some important results. Thus starting with X whose components are X_{rs} we derive $\overline{X}$ whose components are $\overline{X}_{rs} \equiv X_{sr}$ and then the difference $X - \overline{X}$ whose components are $X_{rs} - \overline{X}_{rs} \equiv X_{rs} - X_{sr}$. This new tensor is alternating and an important example of this type will be given to exemplify the next rule.

3. (c) *The Rule of the Simple Product.*

Consider any two tensors not necessarily of the same kind or rank. Let us form the product of each component of the first into each component of the second and arrange the products in a definite order. The set of products will form a tensor whose rank is the sum of the ranks of the original tensors. Again it will suffice to show how the proof runs in a special example. Let the two tensors be X_{rs} and X^{rs} and denote by the symbol $X_{r_1r_2}^{s_1s_2}$ the product $X_{r_1r_2}\cdot X^{s_1s_2}$. (Here r_1, r_2, s_1, s_2 have definite numerical values so that $X_{r_1r_2}^{s_1s_2}$, defined in this way, is a single function out of a group of n^4 obtained by giving r_1, r_2, s_1, s_2 each all values from 1 to n in turn.) We have to show that the group of n^4 functions $X_{r_1r_2}^{s_1s_2}$ really form, as the notation implies, a tensor of rank four covariant of rank two and contravariant of rank two. To do this we have

$$
\begin{aligned}
Y_{r_1r_2}^{s_1s_2} &\equiv Y_{r_1r_2}\cdot Y^{s_1s_2} && \text{by definition of } Y_{r_1r_2}^{s_1s_2}\\
&\equiv \left(X_{\rho_1\rho_2}\frac{\partial x^{(\rho_1)}}{\partial y^{(r_1)}}\frac{\partial x^{(\rho_2)}}{\partial y^{(r_2)}}\right)\cdot\left(X^{\sigma_1\sigma_2}\frac{\partial y^{(s_1)}}{\partial x^{(\sigma_1)}}\cdot\frac{\partial y^{(s_2)}}{\partial x^{(\sigma_2)}}\right) && \\
& && (\rho_1,\ \rho_2,\ \sigma_1,\ \sigma_2 \text{ umbral})\\
&\equiv (X_{\rho_1\rho_2}X^{\sigma_1\sigma_2})\frac{\partial x^{(\rho_1)}}{\partial y^{(r_1)}}\frac{\partial x^{(\rho_2)}}{\partial y^{(r_2)}}\frac{\partial y^{(s_1)}}{\partial x^{(\sigma_1)}}\frac{\partial y^{(s_2)}}{\partial x^{(\sigma_2)}}\\
&\equiv (X_{\rho_1\rho_2}^{\sigma_1\sigma_2})\frac{\partial x^{(\rho_1)}}{\partial y^{(r_1)}}\frac{\partial x^{(\rho_2)}}{\partial y^{(r_2)}}\frac{\partial y^{(s_1)}}{\partial x^{(\sigma_1)}}\frac{\partial y^{(s_2)}}{\partial x^{(\sigma_2)}}\\
& && \text{by definition of } X_{\rho_1\rho_2}^{\sigma_1\sigma_2}
\end{aligned}
$$

which proves the statement.

It is quite apparent that $X_{r_1r_2}^{s_1s_2}$ is not the same as $X_{s_1s_2}^{r_1r_2}$ so that the order of the factors in this kind of a product is important. *Multiplication of tensors is not in general commutative.* This remains true even when both the factors are of the same kind and rank. Consider the simplest case where we have two tensors X and $\bar{X}$ both covariant of rank one. Then the product $X\cdot\bar{X}$ is a tensor $X_{rs}\equiv X_r\cdot\bar{X}_s$ covariant of rank two whilst the product $\bar{X}\cdot X$ is a tensor $\bar{X}_{rs}\equiv\bar{X}_r\cdot X_s$.

The difference $X_{rs} - \overline{X}_{rs}$ is again a covariant tensor of rank wo which is alternating since $\overline{X}_{rs} = X_{sr}$. Since alternating ensors have a more immediate physical significance than non-lternating tensors it is natural to expect that this difference hould be more important than either of the direct products X_{rs} or $\overline{X}_{rs}$. It is what Grassmann called the *outer product* of he two tensors $X, \overline{X}$ in contrast to another kind of product which e calls "*inner*" and which we now proceed to discuss.

. (*d*) ***The Rule of Composition or Inner Multiplication.***

Let us first consider a simple mixed tensor of rank two $X_{r_1}{}^{r_2}$ or which the equations of correspondence are

$$Y_{r_1}{}^{r_2} \equiv X_{s_1}{}^{s_2} \frac{\partial x^{(s_1)}}{\partial y^{(r_1)}} \frac{\partial y^{(r_2)}}{\partial x^{(s_2)}} \quad (s_1 \text{ and } s_2 \text{ umbral symbols})$$

f now we make $r_2 = r_1 = r$ (say) and use r as an umbral symbol e get

$$Y_r{}^r \equiv X_{s_1}{}^{s_2} \frac{\partial x^{(s_1)}}{\partial y^{(r)}} \frac{\partial y^{(r)}}{\partial x^{(s_2)}} \equiv X_{s_1}{}^{s_1} \quad (s_1 \text{ umbral})$$

he remarkable simplification on the right hand side is due to he results from composite differentiation

$$\frac{\partial x^{(s_1)}}{\partial y^{(r)}} \frac{\partial y^{(r)}}{\partial x^{(s_2)}} \equiv \frac{\partial x^{(s_1)}}{\partial x^{(s_2)}}$$
$$\equiv 0 \text{ if } s_2 \neq s_1 \text{ and } \equiv 1 \text{ if } s_2 = s_1$$

n this way we can form from a given tensor a tensor of *lower* ank (in this case an invariant).

The proof in the general case is of the same character.

Consider the mixed tensor $X_{t_1 \dots t_p\, m_1 \dots m_l}^{r_1 \dots r_p\, s_1 \dots s_q}$ which is, as the bels indicate, covariant of rank $p + l$ and contravariant of ank $p + q$ so that the equations of correspondence are

$$Y_{\tau_1 \dots \tau_p\, \mu_1 \dots \mu_l}^{\rho_1 \dots \rho_p\, \sigma_1 \dots \sigma_q} \equiv X_{t_1 \dots t_p\, m_1 \dots m_l}^{r_1 \dots r_p\, s_1 \dots s_q} \frac{\partial y^{(\rho)}}{\partial x^{(r)}} \cdot \frac{\partial y^{(\sigma)}}{\partial x^{(s)}} \cdot \frac{\partial x^{(t)}}{\partial y^{(\tau)}} \cdot \frac{\partial x^{(m)}}{\partial y^{(\mu)}}$$

here $\dfrac{\partial y^{(\rho)}}{\partial x^{(r)}}$ stands for $\dfrac{\partial y^{(\rho_1)}}{\partial x^{(r_1)}} \cdots \dfrac{\partial y^{(\rho_p)}}{\partial x^{(r_p)}}$ and so for the others.

If now we make $\rho_1 = \tau_1, \rho_2 = \tau_2 \cdots \rho_p = \tau_p$ and use $\tau_1 \cdots$ as umbral symbols of summation, $\frac{\partial y^{(\rho)}}{\partial x^{(r)}} \cdot \frac{\partial x^{(t)}}{\partial y^{(\tau)}}$ on the right har side becomes

$$\frac{\partial x^{(t)}}{\partial y^{(\tau)}} \cdot \frac{\partial y^{(\tau)}}{\partial x^{(r)}} \qquad (\tau_1 \cdots \tau_p \text{ umbra}$$

and successive applications of the results

$$\begin{aligned} \frac{\partial x^{(t_1)}}{\partial y^{(\tau_1)}} \frac{\partial y^{(\tau_1)}}{\partial x^{(r_1)}} &= 0 \text{ unless } t_1 = r_1 \\ &= 1 \text{ if } t_1 = r_1 \end{aligned}$$

gives us that

$$\begin{aligned} \frac{\partial x^{(t)}}{\partial y^{(\tau)}} \cdot \frac{\partial y^{(\tau)}}{\partial x^{(r)}} &= 0 \quad \text{unless } t_1 = r_1;\ t_2 = r_2 \cdots t_p = r_p \\ &= 1 \quad \text{if } t_1 = r_1, \cdots, t_p = r_p \end{aligned}$$

so that

$$Y^{\tau_1 \cdots \tau_p \, \sigma_1 \cdots \sigma_q}_{\tau_1 \cdots \tau_p \, \mu_1 \cdots \mu_l} \equiv (X^{r_1 \cdots r_p \, s_1 \cdots s_q}_{r_1 \cdots r_p \, m_1 \cdots m_l}) \frac{\partial y^{(\sigma)}}{\partial x^{(s)}} \cdot \frac{\partial x^{(m)}}{\partial y^{(\mu)}}$$

$$(r, m, s \text{ all umbra}$$

giving the result that $(X^{r_1 \cdots r_p \, s_1 \cdots s_q}_{r_1 \cdots r_p \, m_1 \cdots m_l})$ is a tensor, covariant rank l and contravariant of rank q. If $q = 0, l = 0$ we have tl result that

$$X^{r_1 \cdots r_p}_{r_1 \cdots r_p} \text{ is an invariant} \quad (r_1 \cdots r_p \text{ umbra}$$

explaining why we regard an invariant as a tensor of *zero* ran If now we have two tensors not both entirely covariant contravariant and take their simple product we have a mix tensor to which we may apply the method here described ai obtain a tensor of lower rank. This is called *composition* inner multiplication of the two tensors. Thus starting wi X_r and $\overline{X}^s$ we obtain $X_s^r \equiv X_r \cdot \overline{X}^s$ and then making $r = s$ (i. picking the n diagonal elements or components of the tensor X of rank two) and summing with respect to s we derive an i variant $X_s \cdot \overline{X}^s$ which is the invariant inner product of the tv

tensors. (To obtain an inner product the tensors must be of *different* character—one covariant, the other contravariant.) Similarly from the two tensors of rank two $X^{r_1r_2}$ and $X_{s_1s_2}$ we first obtain the mixed tensor of rank 4

$$X^{r_1r_2}_{s_1s_2} \equiv X^{r_1r_2} \cdot X_{s_1s_2}$$

and from this the scalar or invariant function of position

$$X^{r_1r_2}_{r_1r_2} \equiv X^{r_1r_2} \cdot X_{r_1r_2} \quad (r_1,\ r_2 \text{ umbral symbols})$$

Notice that in these cases the order of the factors is not important—the same invariant results if we change the order.

5. (*e*) *Converse of Rule of Composition.*

Again, for the sake of simplicity, let us explain this for a special case. We consider a set of n functions of position X_r which has such a law of correspondence between components in different coordinate systems that for *any* contravariant tensor $\overline{X}^r$ of rank one *whatsoever* the summation $X_r\overline{X}^r$ is invariant (r umbral). Then we shall prove that the set X_r actually form, as the notation implies, a covariant tensor of rank one.

We have

$$Y_r \cdot \overline{Y}^r \equiv X_t \cdot \overline{X}^{(t)} \qquad \text{(by hypothesis)}$$

$$\equiv X_t \cdot \overline{Y}^s \frac{\partial x^{(t)}}{\partial y^{(s)}} \qquad (\text{since } \overline{X}^r \text{ is contravariant of rank one})$$

We now take as a special illustration of the tensor $\overline{X}^r$ that one, which, in the y system of coordinates, has all its components $\equiv 0$ save one which is $\equiv 1$, e.g., $\overline{Y}^s \equiv 0$ if $s \neq r$ whilst $\overline{Y}^r \equiv 1$. This choice of $\overline{X}$ is permissible since we make the hypothesis that $\overline{X}$ is *any* tensor we wish to choose. And we have

$$Y_r \equiv X_t \frac{\partial x^{(t)}}{\partial y^{(r)}} \qquad (t \text{ umbral})$$

proving on assigning, in turn, to the label r the numerical values $1, \cdots, n$, the statement made. (It is apparent that instead of

taking $\overline{Y}^r$ as perfectly arbitrary it is the same thing to say that $\overline{Y}^{(r)}$ shall be any one of the n tensors which in some particular system of coordinates have each all but one of their coordinates $= 0$, the remaining one being $= 1$.) As another example of this converse let us suppose that the n^2 functions $X_r{}^s$ have such a law of transformation that the summation $X_r{}^s \cdot \overline{X}_{st}$ is a covariant tensor of rank two (s umbral) where $\overline{X}_{st}$ is an arbitrary covariant tensor of rank two; we have to prove that the n^2 functions of position $X_r{}^s$ actually form, as the notation implies, a mixed tensor contravariant of rank 1 and covariant of rank 1.

We have

$$Y_r{}^s\overline{Y}_{st} \equiv (X_\rho{}^\sigma\overline{X}_{\sigma\tau})\frac{\partial x^{(\rho)}}{\partial y^{(r)}}\frac{\partial x^{(\tau)}}{\partial y^{(t)}} \qquad \text{by hypothesis}$$

$$\equiv X_\rho{}^\sigma\overline{Y}_{lm}\frac{\partial y^{(l)}}{\partial x^{(\sigma)}}\cdot\frac{\partial y^{(m)}}{\partial x^{(\tau)}}\frac{\partial x^{(\rho)}}{\partial y^{(r)}}\frac{\partial x^{(\tau)}}{\partial y^{(t)}}$$

(since $\overline{X}$ is covariant of rank 2)

Now as our arbitrary tensor $\overline{X}$ let us choose that one for which

$$\overline{Y}_{lm} \equiv 0 \qquad \text{unless both } l = s \text{ and } m = t$$

$$\overline{Y}_{st} = 1 \qquad \text{and using } \frac{\partial y^{(t)}}{\partial x^{(\tau)}}\cdot\frac{\partial x^{(\tau)}}{\partial y^{(t)}} = 1 \qquad (\tau \text{ umbral})$$

we obtain

$$Y_r{}^s \equiv X_\rho{}^\sigma\frac{\partial y^{(s)}}{\partial x^{(\sigma)}}\frac{\partial x^{(\rho)}}{\partial y^{(r)}} \qquad (\sigma,\ \rho \text{ umbral})$$

proving the statement. The essence of the proof is that the multiplying tensor must be an arbitrary one. In concluding these remarks on the elementary rules of tensor algebra it may not be superfluous to remark that although, for example, the product $X_{rs} \equiv X_r{}^t\cdot\overline{X}_{st}$ is a definite tensor we do not introduce the idea of quotient $X_{rs} \div X_r{}^t$. The reason for this is, of course, that there is no unique quotient; there are many tensors $\overline{X}_{st}$ which when multiplied by a given tensor $X_r{}^t$ in this way will yield a given tensor X_{rs}. In the algebra of tensors it is possible to have a product (inner) of two non-zero tensors equal to zero.

6. *Applications of the Four Rules of Tensor Algebra.*

The most useful applications of these rules will be found by returning to a consideration of the integrals which served to introduce us to the tensor idea. It will be remembered that a curve V_1 is either *open* and has two end points as boundary or else is *closed* and has no boundaries; a spread V_2 of two dimensions is either *open* and bounded by one or more closed curves or *closed* and without boundaries. In general a spread V_{p+1} of $p + 1$ dimensions ($p \leq n - 1$) is either open and bounded by one or more closed spreads V_p of p dimensions or else closed and without boundaries. When the spread V_{p+1} is open there is an important theorem giving the value of an arbitrary integral I_p extended over the closed boundaries V_p in terms of a certain connected integral extended over the open V_{p+1} bounded by V_p. The simplest case is when $p = 1$ in which case an integral over a closed curve is shown to be equivalent to a certain integral extended over any surface or spread of two dimensions V_2 bounded by the curve V_1. This case was discussed by Stokes for ordinary space of 3 dimensions and the general theorem is known as "Stokes' generalized Lemma."* It will be noticed that the theorem is a non-metrical one as we have not yet had occasion to say anything about the metrical character of the space S_n containing the spreads V_p. We shall prove the theorem when $p = 2$ as this will suffice to show the details in the general case.

Here the equations of the open V_3 are

$$x^{(s)} \equiv x^{(s)} (u_1, u_2, u_3) \qquad (s = 1, \cdots, n)$$

and the boundaries will be specified by one or more relations on the parameters u_1, u_2, u_3. If there are several distinct boundaries V_2 we may connect them by auxiliary surfaces V_2 so as to form one complete boundary. The parts of the I_2 over this complete boundary coming from the auxiliary surfaces will cancel (each

* *H. Poincaré*, loc. cit.

auxiliary connecting surface may be replaced by two, infinitesimally close, surfaces and it is the integrals over these pairs of surfaces that cancel each other in the limit as the surfaces are made to approach each other indefinitely). The relation between the parameters on the boundary may be

$$v_3 \equiv \phi(u_1, u_2, u_3) = 0$$

and we introduce two other functions v_1 and v_2 of u_1, u_2, u_3 such that v_1, v_2, v_3 are distinct functions, and change over to v_1, v_2, v_3 as parameters. We shall suppose the parameters such that the equations giving the coordinates x are uniform both ways. Not only does an assigned set of parameters give a unique point x but to a point x there corresponds but one set of parameters v.

Accordingly the surfaces v_3 = const. cannot intersect each other and they form a set of *closed* level surfaces filling up the initial open V_3. On each of these closed level surfaces we shall have the level curves v_1 = const., v_2 = const., and we suppose the functions v_1, v_2 of u_1, u_2, u_3 so chosen that these level curves are *closed*.

Now consider the integral

$$I_2 \equiv \int X_{s_1s_2} d(x^{(s_1)}, x^{(s_2)}) \qquad (s_1, s_2 \text{ umbral and } s_1 < s_2)$$

extended over the boundary $v_3 = 0$. If, instead of integrating over $v_3 = 0$, we take it over any of the level surfaces v_3 = constant it will take on different values depending on this constant and to indicate this we write

$$I_2(v_3) \equiv \int X_{s_1s_2} d(x^{(s_1)} x^{(s_2)}) \qquad (s_1 < s_2)$$

$$\equiv \int \left(X_{s_1s_2} \frac{\partial x^{(s_1)}}{\partial v_1} \frac{\partial x^{(s_2)}}{\partial v_2} \right) dv_1 dv_2$$

$$\frac{dI_2}{dv_3} = \int \left\{ X_{s_1s_2} \left(\frac{\partial^2 x^{(s_1)}}{\partial v_1 \partial v_3} \frac{\partial x^{(s_2)}}{\partial v_2} + \frac{\partial x^{(s_1)}}{\partial v_1} \frac{\partial^2 x^{(s_2)}}{\partial v_2 \partial v_3} \right) + \frac{\partial X_{s_1s_2}}{\partial v_3} \frac{\partial x^{(s_1)}}{\partial v_1} \frac{\partial x^{(s_2)}}{\partial v_2} \right\} dv_1 dv_2$$

(It is only necessary to differentiate the integrand since the limits of the integral are independent of v_3). Now if F is any function of *position* (not merely of the parameters)* on a closed curve with parameter v the integral $\int \frac{\partial F}{\partial v} dv$ taken round the *closed* curve is necessarily zero. For it is the difference of the values of F at the coincident end points of the curve. If, in particular, we take as F the function

$$F \equiv X_{s_1 s_2} \frac{\partial x^{(s_1)}}{\partial v_3} \frac{\partial x^{(s_2)}}{\partial v_2} \qquad (s_1, s_2 \text{ umbral})$$

and integrate round the closed curve v_2 = constant we get

$$\int \left\{ X_{s_1 s_2} \left(\frac{\partial^2 x^{(s_1)}}{\partial v_1 \partial v_3} \frac{\partial x^{(s_2)}}{\partial v_2} + \frac{\partial x^{(s_1)}}{\partial v_3} \frac{\partial^2 x^{(s_2)}}{\partial v_1 \partial v_2} \right) + \frac{\partial X_{s_1 s_2}}{\partial v_1} \frac{\partial x^{(s_1)}}{\partial v_3} \frac{\partial x^{(s_2)}}{\partial v_2} \right\} dv_1 \equiv 0$$

and integrating this with respect to v_2 over the surface v_3 = constant we have

$$\int \left\{ X_{s_1 s_2} \left(\frac{\partial^2 x^{(s_1)}}{\partial v_1 \partial v_3} \frac{\partial x^{(s_2)}}{\partial v_2} + \frac{\partial x^{(s_1)}}{\partial v_3} \frac{\partial^2 x^{(s_2)}}{\partial v_1 \partial v_2} \right) + \frac{\partial X_{s_1 s_2}}{\partial v_1} \frac{\partial x^{(s_1)}}{\partial v_3} \frac{\partial x^{(s_2)}}{\partial v_2} \right\} dv_1 dv_2 \equiv 0$$

Similarly on taking

$$F \equiv X_{s_1 s_2} \frac{\partial x^{(s_2)}}{\partial v_3} \frac{\partial x^{(s_1)}}{\partial v_1}$$

and integrating $\int \frac{\partial F}{\partial v_2} dv_1 dv_2$ over the closed surface v_3 = const.

* The distinction implied here should be clearly grasped. If the equations of the curve are

$$x_1 = a \cos v$$
$$x_2 = a \sin v$$

F must be periodic in v with period 2π.

we get

$$\int\left\{X_{s_1s_2}\left(\frac{\partial^2x^{(s_2)}}{\partial v_2\partial v_3}\frac{\partial x^{(s_1)}}{\partial v_1}+\frac{\partial x^{(s_2)}}{\partial v_3}\frac{\partial^2x^{(s_1)}}{\partial v_1\partial v_2}\right)\right.$$
$$\left.+\frac{\partial X_{s_1s_2}}{\partial v_2}\frac{\partial x^{(s_2)}}{\partial v_3}\frac{\partial x^{(s_1)}}{\partial v_1}\right\}dv_1dv_2\equiv 0$$

Now add these two equations together and note that

$$X_{s_1s_2}\left(\frac{\partial x^{(s_1)}}{\partial v_3}\frac{\partial^2x^{(s_2)}}{\partial v_1\partial v_2}+\frac{\partial x^{(s_2)}}{\partial v_3}\frac{\partial^2x^{(s_1)}}{\partial v_1\partial v_2}\right)\equiv 0\qquad (s_1,\,s_2\text{ umbral})$$

because the terms in the summation cancel out in pairs owing to the alternating character of $X_{s_1s_2}$—the factor multiplying $X_{s_1s_2}$ in the summation being obviously unaltered by an interchange of the symbols s_1 and s_2. We find that

$$\int\left\{X_{s_1s_2}\left(\frac{\partial^2x^{(s_1)}}{\partial v_1\partial v_3}\frac{\partial x^{(s_2)}}{\partial v_2}+\frac{\partial^2x^{(s_2)}}{\partial v_2\partial v_3}\frac{\partial x^{(s_1)}}{\partial v_1}\right)\right.$$
$$\left.+\frac{\partial X_{s_1s_2}}{\partial v_1}\frac{\partial x^{(s_1)}}{\partial v_3}\frac{\partial x^{(s_2)}}{\partial v_2}+\frac{\partial X_{s_1s_2}}{\partial v_2}\frac{\partial x^{(s_2)}}{\partial v_3}\frac{\partial x^{(s_1)}}{\partial v_1}\right\}dv_1dv_2\equiv 0$$

so that

$$\frac{dI_2}{dv_3}\equiv\int\left\{\frac{\partial X_{s_1s_2}}{\partial v_3}\frac{\partial x^{(s_1)}}{\partial v_1}\frac{\partial x^{(s_2)}}{\partial v_2}-\frac{\partial X_{s_1s_2}}{\partial v_1}\frac{\partial x^{(s_1)}}{\partial v_3}\frac{\partial x^{(s_2)}}{\partial v_2}\right.$$
$$\left.-\frac{\partial X_{s_1s_2}}{\partial v_2}\frac{\partial x^{(s_2)}}{\partial v_3}\frac{\partial x^{(s_1)}}{\partial v_1}\right\}dv_1dv_2$$

Now the $X_{s_1s_2}$ are functions of position, i.e., of the coordinates x so that

$$\frac{\partial X_{s_1s_2}}{\partial v_3}\equiv\frac{\partial X_{s_1s_2}}{\partial x^{(s_3)}}\frac{\partial x^{(s_3)}}{\partial v_3}\qquad (s_3\text{ umbral})$$

The second term in dI_2/dv_3 we shall slightly modify by a change in the umbral symbols. Thus

$$\frac{\partial X_{s_1s_2}}{\partial v_1}\frac{\partial x^{(s_1)}}{\partial v_3}\frac{\partial x^{(s_2)}}{\partial v_2}\equiv\frac{\partial X_{s_1s_2}}{\partial x^{(s_3)}}\frac{\partial x^{(s_3)}}{\partial v_1}\frac{\partial x^{(s_1)}}{\partial v_3}\frac{\partial x^{(s_2)}}{\partial v_2}\qquad (s_1,\,s_2,\,s_3\text{ all umbral})$$
$$\equiv\frac{\partial X_{s_3s_2}}{\partial x^{(s_1)}}\frac{\partial x^{(s_1)}}{\partial v_1}\frac{\partial x^{(s_2)}}{\partial v_2}\frac{\partial x^{(s_3)}}{\partial v_3}$$

so that we can write

$$\frac{dI_2}{dv_3} \equiv \int \left\{ \frac{\partial X_{s_1 s_2}}{\partial x^{(s_3)}} - \frac{\partial X_{s_3 s_2}}{\partial x^{(s_1)}} - \frac{\partial X_{s_1 s_3}}{\partial x^{(s_2)}} \right\} \frac{\partial x^{(s_1)}}{\partial v_1} \frac{\partial x^{(s_2)}}{\partial v_2} \frac{\partial x^{(s_3)}}{\partial v_3} dv_1 dv_2$$

On writing

$$X_{s_1 s_2 s_3} \equiv \frac{\partial X_{s_1 s_2}}{\partial x^{(s_3)}} - \frac{\partial X_{s_3 s_2}}{\partial x^{(s_1)}} - \frac{\partial X_{s_1 s_3}}{\partial x^{(s_2)}}$$

and integrating the expression for dI_2/dv_3 with respect to v_3 we find

$$I_2 \Big| \equiv \int X_{s_1 s_2 s_3} \frac{\partial x^{(s_1)}}{\partial v_1} \frac{\partial x^{(s_2)}}{\partial v_2} \frac{\partial x^{(s_3)}}{\partial v_3} dv_1 dv_2 dv_3 \qquad (s_1, s_2, s_3 \text{ umbral})$$

$$\equiv \int X_{s_1 s_2 s_3} d(x^{(s_1)}, x^{(s_2)}, x^{(s_3)}) \qquad (s_1 < s_2 < s_3)$$

since the set of functions $X_{s_1 s_2 s_3}$ defined as above is obviously alternating (on account of the fact that X_{rs} is an alternating set). The limits for v_3 are $v_3 = 0$ and $v_3 =$ some constant for which $I_2 = 0$—since the corresponding V_2 is either a point or a spread traced twice on opposite sides. Let the integration be such that $v_3 = 0$ is the *upper* limit and we have

$$I_2 \equiv \int X_{s_1 s_2} d(x^{(s_1)} x^{(s_2)}) \qquad (s_1 < s_2) \text{ over boundary}$$
$$\equiv \int X_{s_1 s_2 s_3} d(x^{(s_1)} x^{(s_2)} x^{(s_3)}) \; (s_1 < s_2 < s_3) \text{ over the } V_3.^*$$

In general from

$$I_p \equiv \int X_{s_1 \cdots s_p} d(x^{(s_1)} \cdots x^{(s_p)}) \quad (s_1 < s_2 \cdots < s_p)$$

over a closed boundary we derive as equivalent to I_p an

$$I_{p+1} \equiv \int X_{s_1 \cdots s_{p+1}} d(x^{(s_1)} \cdots x^{(s_{p+1})}) \quad (s_1 < s_2 \cdots < s_{p+1})$$

where

$$X_{s_1 \cdots s_{p+1}} \equiv \frac{\partial X_{s_1 \cdots s_p}}{\partial x^{(s_{p+1})}} - \frac{\partial X_{s_{p+1} s_2 \cdots s_p}}{\partial x^{(s_1)}} - \frac{\partial X_{s_1 s_{p+1} s_3 \cdots s_p}}{\partial x^{(s_2)}} \cdots - \frac{\partial X_{s_1 \cdots s_{p-1} s_{p+1}}}{\partial x^{(s_p)}}$$

* It will be observed that placing the + sign before I_2 on the left makes $v_3 = 0$ the *upper* bound of the integral $\int \frac{dI_2}{dv_3} dv_3$. Thus v_3 is increasing away from the open spread V_3.

It is usual to preserve a cyclic arrangement of suffixes for the X's and then, on account of the alternating character of the X's, we have

$$X_{s_1 \cdots s_{p+1}} = \frac{\partial X_{s_1 \cdots s_p}}{\partial x^{(s_{p+1})}} \pm \frac{\partial X_{s_2 \cdots s_p s_{p+1}}}{\partial x^{(s_1)}} + \frac{\partial X_{s_3 \cdots s_{p+1} s_1}}{\partial x^{(s_2)}} \pm \cdots$$

the upper signs being used when p is even and the lower when p is odd. Since I_p is by hypothesis invariant so is I_{p+1} because $I_{p+1} = I_p$ and accordingly the coefficients $X_{s_1 \cdots s_{p+1}}$ form an alternating covariant tensor of rank $p + 1$ [seen either directly as when tensors were introduced or as a case of the converse of rule (d), the set of functions $\frac{\partial x^{(s_1)}}{\partial v_1} \cdots \frac{\partial x^{(s_{p+1})}}{\partial v_{p+1}} dv_1 \cdots dv_{p+1}$ forming an arbitrary contravariant tensor of rank $p + 1$]. In this way we can derive from any alternating covariant tensor, by a species of differentiation, a covariant tensor of higher rank.

EXAMPLES.

$p = 1$. From any covariant tensor X_r of rank one we derive an alternating covariant tensor of rank two

$$X_{rs} \equiv \frac{\partial X_r}{\partial x^{(s)}} - \frac{\partial X_s}{\partial x^{(r)}}$$

It is the negative of this tensor that is called the *curl of the vector* X in the earlier vector analysis. It is rather important to notice that this, and the other tensors of this paragraph, have no reference to the metrical character of the fundamental space S_n. The derivation of them by the methods of the Absolute Differential Calculus introduces, therefore, extraneous and unnecessary ideas.

$p = 2$. From an alternating covariant tensor of rank two X_{rs} we derive the alternating covariant tensor of rank three

$$X_{rst} \equiv \frac{\partial X_{rs}}{\partial x^{(t)}} + \frac{\partial X_{st}}{\partial x^{(r)}} + \frac{\partial X_{tr}}{\partial x^{(s)}}$$

If $n = 3$ there is only one such function and in th
it is called the divergence of X_{rs}. We shall hav
slightly for the general tensor analysis. It is
notice that if we take as X_{rs} the tensor of the I

$$X_{rs} \equiv \frac{\partial X_r}{\partial x^{(s)}} - \frac{\partial X_s}{\partial x^{(r)}}$$

we find $X_{rst} \equiv 0$. It is easily seen that this happens in general. If we derive $X_{s_1 \dots s_p}$ from $X_{s_1 \dots s_{p-1}}$ in this way then the $X_{s_1 \dots s_{p+1}}$ derived from $X_{s_1 \dots s_p}$ is $\equiv 0$. When the $X_{s_1 \dots s_{p+1}}$ derived from $X_{s_1 \dots s_p}$ is $\equiv 0$ we have that $I_{p+1} \equiv 0$ and so I_p (extended, of course, over any *closed* spread of p dimensions) is $\equiv 0$. In this case I_p is said to be the *integral of an exact differential*. It can then be proved that the value of I_p over any *open* V_p is equal to the value of a certain integral I_{p-1} over the closed boundary of this V_p.*

* If

$$I_p \equiv \int X_{s_1 s_2 \dots s_p} d(x^{(s_1)} \cdots x^{(s_p)}) \qquad (s_1 < s_2 \cdots < s_p)$$

is the integral of an exact differential we have $\binom{n}{p+1} \equiv \dfrac{n!}{\overline{n-p-1}!\,\overline{p+1}!}$ partial differential equations

$$X_{s_1 s_2 \dots s_{p+1}} \equiv 0$$

The theorem stated is that these are the necessary and sufficient conditions that there exist $\binom{n}{p-1}$ functions of position $X_{s_1 \dots s_{p-1}}$ satisfying the $\binom{n}{p}$ partial differential equations

$$\frac{\partial X_{s_1 \dots s_{p-1}}}{\partial x^{(s_p)}} - \frac{\partial X_{s_p s_2 \dots s_{p-1}}}{\partial x^{(s_1)}} - \frac{\partial X_{s_1 s_p s_3 \dots s_{p-1}}}{\partial x^{(s_2)}} \cdots - \frac{\partial X_{s_1 s_2 \dots s_{p-2} s_p}}{\partial x^{(s_{p-1})}} \equiv X_{s_1 s_2 \dots s_p}$$

That the conditions are necessary is an immediate result of a direct substitution of the left hand side of the equation just written for $X_{s_1 \dots s_p}$ in the equation of definition

$$X_{s_1 \dots s_{p+1}} \equiv \frac{\partial X_{s_1 \dots s_p}}{\partial x^{(s_{p+1})}} - \frac{\partial X_{s_{p+1} s_2 \dots s_p}}{\partial x^{(s_1)}} - \dots - \frac{\partial X_{s_1 s_2 \dots s_{p-1} s_{p+1}}}{\partial x^{(s_p)}}$$

To prove the sufficiency an appeal is made to the principle of mathematical induction. Let us, for definiteness, take $p = 2$. Then we shall prove the statement that if the theorem is true for a particular value of n it is true for the next greater integer value $n + 1$. Granting this, for the moment, we

$p = n - 1$. This is the next and last case if $n = 4$. For an arbitrary value of n it is second in importance only to the first case $p = 1$. In order to avoid having to write out separately

observe that the theorem is true for $n = 2$. (In this case there are no integrability conditions necessary; on account of the alternating character of the Tensor X_{qrs} whose vanishing expresses these conditions, it is necessarily $= 0$.) We have two unknowns X_1 and X_2 satisfying the single differential equation

$$\frac{\partial X_1}{\partial x^{(2)}} - \frac{\partial X_2}{\partial x^{(1)}} = X_{12}$$

and a particular solution is found by assuming that neither X_1 nor X_2 involves $x^{(2)}$. Then X_1 may be any function of $x^{(1)}$ and $X_2 = -\int^{x^{(1)}} X_{12} dx^{(1)}$, the lower limit being any *constant* $x_0^{(1)}$. In the integration $x^{(2)}$ is regarded as a constant. Hence by the induction lemma the theorem is true for $n = 3$ and then for $n = 4$ and so for every integer n.

To prove the induction lemma let us seek for a solution of the equations

$$X_{rs} = \frac{\partial X_r}{\partial x^{(s)}} - \frac{\partial X_s}{\partial x^{(r)}} \qquad (r < s, = 1, \cdots, n)$$

unknown $X_n = 0$. We have then

$$X_{rn} = + \frac{\partial X_r}{\partial x^{(n)}} \qquad (r = 1, \cdots, n-1)$$

whence

$$X_r = + \int_{x_0^{(n)}}^{x^{(n)}} X_{rn} dx^{(n)} + \overline{X}_r \qquad (r = 1, \cdots, n-1)$$

where $x_0^{(n)}$ is a constant; $\overline{X}_r$ is any function of $x^{(1)}, \cdots, x^{(n-1)}$ and in the integration $x^{(1)}, \cdots, x^{(n-1)}$ are constants. The remaining equations

$$X_{rs} = \frac{\partial X_r}{\partial x^{(s)}} - \frac{\partial X_s}{\partial x^{(r)}} \qquad (r < s = 1, \cdots, n-1)$$

give on substituting these values

$$X_{rs} = + \int_{x_0^{(n)}}^{x^{(n)}} \frac{\partial X_{rn}}{\partial x^{(s)}} dx^{(n)} - \int_{x_0^{(n)}}^{x^{(n)}} \frac{\partial X_{sn}}{\partial x^{(r)}} dx^{(n)} + \frac{\partial \overline{X}_r}{\partial x^{(s)}} - \frac{\partial \overline{X}_s}{\partial x^{(n)}}$$

$$= + \int_{x_0^{(n)}}^{x^{(n)}} \frac{\partial X_{rs}}{\partial x^{(n)}} dx^{(n)} + \frac{\partial \overline{X}_r}{\partial x^{(s)}} - \frac{\partial \overline{X}_s}{\partial x^{(r)}}$$

from $0 = X_{rns} = \dfrac{\partial X_{rn}}{\partial x^{(s)}} + \dfrac{\partial X_{ns}}{\partial x^{(r)}} + \dfrac{\partial X_{sr}}{\partial x^{(n)}}$

$= X_{rs} - \overline{X}_{rs} + \dfrac{\partial \overline{X}_r}{\partial x^{(s)}} - \dfrac{\partial \overline{X}_s}{\partial x^{(r)}}$ where $\overline{X}_{rs}$ is the function X_{rs} when $x^{(n)}$ is put $= x_0^{(n)}$.

the cases corresponding to n even and n odd we shall adopt the first form for $X_{s_1 \cdots s_{p+1}}$.

$$X_{s_1 \cdots s_{p+1}} \equiv \frac{\partial X_{s_1 \cdots s_p}}{\partial x^{(s_{p+1})}} - \frac{\partial X_{s_{p+1}s_2 \cdots s_{p-1}s_p}}{\partial x^{(s_1)}} - \cdots - \frac{\partial X_{s_1 \cdots s_{p-1}s_{p+1}}}{\partial x^{(s_p)}}$$

Hence we have the $\binom{n-1}{2}$ equations

$$\overline{X}_{rs} = \frac{\partial \overline{X}_r}{\partial x^{(s)}} - \frac{\partial \overline{X}_s}{\partial x^{(r)}}$$

with $n-1$ unknowns $\overline{X}_r$ and involving $n-1$ independent variables $x^{(1)}, \cdots, x^{(n-1)}$. Also we have $\binom{n-1}{3}$ integrability equations $\overline{X}_{rst} = 0$ found by putting $x^{(n)} = x_0^{(n)}$ in

$$X_{rst} = 0 \qquad (r < s < t = 1, \cdots, n-1)$$

Hence if we can solve these equations, i.e., if our hypothesis is true for $n-1$, we can solve the original equations which are identical in form but involve one more independent variable $x^{(n)}$. The particular case of this theorem corresponding to $n = 4$, $p = 2$, tells us that Maxwell's equations

$$\operatorname{curl} \overline{E} + \frac{1}{c}\frac{\partial \overline{B}}{\partial t} = 0 \qquad \operatorname{div} \overline{B} = 0 \quad \text{(in the usual notation)}$$

imply the existence of the electromagnetic potential $(A_x, A_y, A_z, -c\phi)$—which is as in the general case when $p = 2$ a covariant tensor of rank one—such that

$$\overline{B} = \operatorname{curl} \overline{A};\ \overline{E} = -\operatorname{grad} \phi - \frac{1}{c}\frac{\partial \overline{A}}{\partial t}$$

For further details cf. *Physical Review*, N. S., Vol. 17, p. 83 (1921).

It is apparent that there is a great degree of arbitrariness allowed in the determination of the functions $X_{s_1 \cdots s_{p-1}}$; in fact we may add to any solution any alternating covariant tensor of rank $p-1$ whose integral over any closed spread V_{p-1} of $p-1$ dimensions is zero. For example we may add to the electromagnetic potential any *gradient of a function of position;* that is if $(A_x, A_y, A_z, -c\phi)$ is any determination of the electromagnetic potential, so is

$$\left.\begin{aligned} \overline{A}_x &\equiv A_x + \frac{\partial F}{\partial x} \\ \overline{A}_y &\equiv A_y + \frac{\partial F}{\partial y} \\ \overline{A}_z &\equiv A_z + \frac{\partial F}{\partial z} \\ \overline{\varphi} &\equiv \varphi - \frac{1}{c}\frac{\partial F}{\partial t} \end{aligned}\right\} \text{where } F \text{ is an arbitrary function of } x, y, z, t.$$

Here $p + 1 = n$ and there is only one distinct function $X_{s_1 \dots s_n}$ on account of the alternating character of this set. Let us choose this one as $X_{1 \dots n}$ and our formula is

$$X_{1 \dots n} = \frac{\partial X_{1 \dots n-1}}{\partial x^{(n)}} - \frac{\partial X_{n2 \dots n-1}}{\partial x^{(1)}} - \dots - \frac{\partial X_{1,\,2 \dots n-2\,n}}{\partial x^{(n-1)}}$$

Now there are only n distinct functions $X_{s_1 \dots s_{n-1}}$ and it will be possible, and convenient, to indicate these by means of a single label. Thus we write

$$\begin{aligned}
(X_n) &= X_{1 \dots n-1} \\
(X_{n-1}) &= -X_{12 \dots n-2,\,n} \\
(X_{n-2}) &= +X_{12 \dots n-3,\,n-1,\,n} \equiv -X_{12 \dots n-3,\,n,\,n-1} \\
&\dots\dots\dots\dots \\
(X_1) &= (-1)^{n-1} X_{23 \dots n} \equiv -X_{n23 \dots n-1}
\end{aligned}$$

where we are careful to put parentheses round the symbols (X_r) to indicate that they are *not* the components of a covariant tensor of rank one.

Maxwell availed himself of this arbitrariness and chose F so that $\operatorname{div} \bar{A} = 0$ whence

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \frac{\partial^2 F}{\partial z^2} = -\operatorname{div} A$$

yielding, from the theory of the Newtonian Potential,

$$F = \frac{1}{4\pi} \int \frac{\operatorname{div} A}{r}\, dt$$

The usual procedure with modern writers is to choose F so that

$$\operatorname{div} \bar{A} + \frac{1}{c}\frac{\partial \bar{\phi}}{\partial t} = 0$$

The equation determining F is now

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + \frac{\partial^2 F}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 F}{\partial t^2} = -\left(\operatorname{div} A - \frac{1}{c}\frac{\partial \phi}{\partial t}\right),$$

whence

$$F = \frac{1}{4\pi} \int \frac{\left(\operatorname{div} A - \frac{1}{c}\frac{\partial \phi}{\partial t}\right)_{t-\frac{r}{c}}}{r}\, d\tau$$

from the theory of the *retarded* potential.

Then we have

$$X_{1 \ldots n} \equiv \frac{\partial(X_s)}{\partial x^{(s)}} \qquad (s \text{ an umbral label})$$

Although the (X_s) do not form a covariant tensor of rank one they are very closely related to a *contravariant* tensor of rank one. In fact there is a reciprocal relationship between an *alternating covariant tensor* of any rank r and an allied *contravariant* alternating tensor of rank $n - r$. It is a special case of this reciprocity stressed so much by Grassmann in his Ausdehnungslehre that gives the dual relationship of point and plane, line and line in analytic projective geometry and it is from the terminology of that subject that the terms " covariant " and " contravariant " are taken. In order to bring out this reciprocal relationship in the clearest manner we must make a digression and discuss what are meant by " metrical properties " of space.

CHAPTER III

1. Introduction of the metrical idea into our geometry*

Let us consider a curve V_1 specified by the equations

$$x^{(s)} \equiv x^{(s)}(u) \qquad (s = 1, \cdots, n)$$

The quadratic differential form

$$g_{rs}dx^{(r)}dx^{(s)} \qquad (r, s \text{ umbral})$$

where the g_{rs} are functions of position, will be invariant provided that these functions form a covariant tensor of rank 2. (This is a consequence of our rule (*d*), Ch. 2, § 4, and its converse since the set of n^2 functions

$$dx^r \cdot dx^s \equiv \frac{dx^{(r)}}{du} \cdot \frac{dx^{(s)}}{du} (du)^2$$

form a contravariant tensor of rank two.) Accordingly the g_{rs} being of this kind the integral

$$s \equiv \int_{u0}^{u'} \sqrt{g_{rs} \frac{dx^{(r)}}{du} \frac{dx^{(s)}}{du}} \cdot du$$

has a value independent of the choice of coordinates x; it is called the length of the curve V_1 from the point specified by u_0 to that specified by u'. If the upper limit u' is regarded as variable and written, therefore, without the prime S is a function of this upper limit u and its differential is given by

$$(ds)^2 = g_{rs}dx^{(r)}dx^{(s)} \qquad (r, s \text{ umbral})$$

where the positive radical is taken on extracting the square root. It will be convenient to agree that, in some particular set of coordinates x, we arrange matters so that $g_{rs} \equiv g_{sr}$; this can always

* The most satisfactory presentation of the general idea of a metrical space is that given in *Bianchi, L.*, Lezioni di Geometria Differenziale, Vol. 1, § 152.

be done by rewriting any two terms, $g_{23}dx^{(2)}dx^{(3)} + g_{32}dx^{(3)}dx^{(2)}$ for example, of the summation which do not satisfy this requirement in the form $\frac{1}{2}(g_{23} + g_{32})dx^{(2)}dx^{(3)} + \frac{1}{2}(g_{23} + g_{32})dx^{(3)}dx^{(2)}$.

The equations defining the covariant correspondence

$$f_{rs} \equiv g_{lm}\frac{\partial x^{(l)}}{\partial y^{(r)}}\frac{\partial x^{(m)}}{\partial y^{(s)}} \qquad (l,\ m \text{ umbral})$$

where

$$(ds)^2 \equiv f_{rs}dy^{(r)}dy^{(s)}$$

then show that

$$\begin{aligned} f_{sr} &\equiv g_{lm}\frac{\partial x^{(l)}}{\partial y^{(s)}}\frac{\partial x^{(m)}}{\partial y^{(r)}} \\ &\equiv g_{ml}\frac{\partial x^{(l)}}{\partial y^{(s)}}\frac{\partial x^{(m)}}{\partial y^{(r)}} \qquad \text{since } g_{rs} \equiv g_{sr} \\ &\equiv f_{rs} \end{aligned}$$

We may express this result by saying that the property of any special tensor of being symmetric is an absolute one just as is the property of being alternating.

2. Reciprocal form for $(ds)^2$

Consider the *n linear differential forms*

$$\xi_r \equiv g_{rs}dx^{(s)} \quad (s \text{ umbral};\ r = 1, \cdots, n)$$

We can solve these for the differentials $dx^{(s)}$ in terms of the n quantities ξ_r as follows. (Note that the ξ_r form, as the notation indicates, a covariant tensor of rank 1 from our rule (d) of composition or inner multiplication.) Let us denote the cofactor of any element g_{rs} in the expansion of the determinant

$$g \equiv \begin{vmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ g_{n1} & \cdots & & g_{nn} \end{vmatrix}$$

by (G_{rs}), observing in passing that $(G_{rs}) \equiv (G_{sr})$. The parentheses indicate that the (G_{rs}) do not form a tensor. From the

definition of a cofactor the summation

$$g_{rs}(G_{rm}) \equiv g \quad \text{when } m = s \qquad (r \text{ umbral})$$
$$\equiv 0 \quad \text{when } m \neq s$$

We shall now introduce the hypothesis that our metrical space is such that g does not vanish identically (it will be presently seen that this is an absolute property) and for all points where g is not zero we have

$$g_{rs}\frac{(G_{rm})}{g} = 1 \qquad \text{when } m = s \qquad (r \text{ umbral})$$
$$= 0 \qquad \text{when } m \neq s$$

Write $g^{lm} \equiv (G_{lm})/g$ and let us justify the notation by showing that the g^{lm} form a contravariant tensor of rank two. From our definition it is symmetrical and so we have in addition to

$$g_{rs}g^{rm} = 1 \qquad \text{if } m = s$$
$$= 0 \qquad \text{if } m \neq s$$

the equivalent equations

$$g_{sr}g^{mr} = 1 \qquad \text{if } m = s$$
$$= 0 \qquad \text{if } m \neq s$$

These relations suggest that we multiply the equations of definition

$$\xi_r \equiv g_{rs}dx^{(s)}$$

by g^{rm} and use r as an umbral symbol. We obtain then

$$g^{rm}\xi_r \equiv g_{rs}g^{rm}dx^{(s)} \qquad (r,\ s \text{ umbral})$$
$$\equiv dx^{(m)} \quad \text{from our relations just written}$$

Accordingly

$$(ds)^2 \equiv g_{lm}dx^{(l)}dx^{(m)} \equiv g_{lm}g^{lr}\xi_r \cdot g^{ms}\xi_s \quad (l,\ m,\ r,\ s \text{ umbral})$$
$$\equiv g^{rs}\xi_r\xi_s \qquad (r,\ s \text{ umbral})$$

since $g_{lm}g^{lr} \equiv 0$ unless $m = r$ when it $\equiv 1$.

The $\xi_r\xi_s$ form, by rule (c), Ch. 2, § 3, an arbitrary contravariant tensor of rank 2 and $(ds)^2$ being, by hypothesis, invariant, the

converse of rule (*d*), Ch. 2, § 5, gives us the result that the g^{rs} form a contravariant (symmetrical) tensor of rank 2. When we write

$$(ds)^2 \equiv g^{rs}\xi_r\xi_s \qquad (r, s \text{ umbral})$$

it is said to be written in the *reciprocal form*. We could start with this form and write

$$dx^{(r)} \equiv g^{rs}\xi_s$$

and solving these obtain

$$\xi_s \equiv g_{sr}dx^{(r)}$$

and then find

$$(ds)^2 \equiv g_{rs}dx^{(r)}dx^{(s)}$$

3. If now we have two determinants $a \equiv |a_{rs}|$, $b \equiv |b_{rs}|$ each of order n (the notation implying that a_{rs} is the element in the rth row and sth column of the determinant a) it is well known that the product of the determinants a and b may be written as a determinant c* of which the elements c_{rs} are defined by

$$c_{rs} \equiv a_{lr}b_{ls} \qquad (l \text{ an umbral symbol})$$

This kind of a product is said to be taken by multiplying columns of a into columns of b.

We can, with the aid of this rule, easily see how the determinant g behaves when we change our coordinates x to some other suitable coordinates y. We get a determinant f of which the r, sth element is

$$f_{rs} \equiv g_{lm}\frac{\partial x^{(l)}}{\partial y^{(r)}}\frac{\partial x^{(m)}}{\partial y^{(s)}}$$

Here $\dfrac{\partial x^{(l)}}{\partial y^{(r)}}$ may be conveniently denoted by (j_{lr}) since it is the l, rth element of the Jacobian determinant J of the transformation from x to y coordinates

* Cf. *Bôcher, M., Introduction to Higher Algebra*, Chap. 2, Macmillan (1915).

$$J = \begin{vmatrix} \dfrac{\partial x^{(1)}}{\partial y^{(1)}} & \cdots & \dfrac{\partial x^{(1)}}{\partial y^{(n)}} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \dfrac{\partial x^{(n)}}{\partial y^{(1)}} & \cdots & \dfrac{\partial x^{(n)}}{\partial y^{(n)}} \end{vmatrix}$$

and then

$$g_{lm}\frac{\partial x^{(l)}}{\partial y^{(r)}} \equiv g_{lm}(j_{lr})$$

is the *mr*th element of the product gJ so that $\left(g_{lm}\dfrac{\partial x^{(l)}}{\partial y^{(r)}}\right)\dfrac{\partial x^{(m)}}{\partial y^{(s)}}$ is the *rs*th element of the product of the determinants gJ by J. Hence $f = gJ^2$.

This important formula shows us that if $g \neq 0$ neither will $f \equiv 0$ unless $J \equiv 0$ in which case the y's would not be suitable coordinates. f can be zero at points where $g \neq 0$ if $J = 0$ at those points; such points would be singular points of the system of coordinates and the quantities f^{rs} would not be defined for them.

Example

In Euclidean space of 3 dimensions with rectangular Cartesian coordinates $x^{(1)}$ $x^{(2)}$ $x^{(3)}$ we write

$$(ds)^2 = (dx^{(1)})^2 + (dx^{(2)})^2 + (dx^{(3)})^2$$

so that $g_{11} = g_{22} = g_{33} = 1$, $g_{12} = g_{13} = g_{23} = 0$. In space polar coordinates we find

$$f_{11} \equiv 1 \qquad f_{22} \equiv (y^{(1)})^2 \qquad f_{33} \equiv y^{(1)^2}\sin^2 y^{(2)}$$
$$f_{12} = f_{13} = f_{23} = 0.$$

Here $g = 1$

$$f = f_{11}f_{22}f_{33} = J^2$$

so that

$$g^{11} = g^{22} = g^{33} = 1 \qquad g^{12} = g^{13} = g^{23} = 0$$

$$f^{11} = \frac{1}{f_{11}} = 1; \quad f^{22} = \frac{1}{f_{22}} = \frac{1}{y^{(1)^2}}; \quad f^{33} = \frac{1}{f_{33}} = \frac{1}{y^{(1)^2}\sin^2 y^{(2)}};$$
$$f^{12} = f^{13} = f^{23} = 0$$

and

$$(ds)^2 = +\xi_1^2 + \xi_2^2 + \xi_3^2.$$

In fact $\xi_1 = dx^{(1)}$, etc. There are no singular points in the x coordinates but there are in the y system; those for which $J = 0$, i.e.,

$$y^{(1)^2} \sin y^{(2)} = 0$$

These are the points on the polar axis

$$y^{(1)} \equiv r = 0; \qquad y^{(2)} \equiv \theta = 0 \text{ or } \pi$$

4. If now $u_1 \cdots u_n$ are any independent parameters in terms of which it is convenient to specify both the x and y coordinates we have, by definition of the symbol,

$$d(y^{(1)} \cdots y^{(n)}) \equiv \frac{\partial(y^{(1)} \cdots y^{(n)})}{\partial(u_1 \cdots u_n)} du_1 \cdots du_n$$

and a similar equation for $d(x^{(1)} \cdots x^{(n)})$ so that

$$\frac{d(y^{(1)} \cdots y^{(n)})}{d(x^{(1)} \cdots x^{(n)})} \equiv \frac{\partial(y^{(1)} \cdots y^{(n)})}{\partial(u_1 \cdots u_n)} \div \frac{\partial(x^{(1)} \cdots x^{(n)})}{\partial(u_1 \cdots u_n)}$$

If we multiply the determinants $\dfrac{\partial(x^{(1)} \cdots x^{(n)})}{\partial(u_1 \cdots u_n)}$ and $\dfrac{\partial(u_1 \cdots u_n)}{\partial(x^{(1)} \cdots x^{(n)})}$ together and note that

$$\begin{aligned} \frac{\partial x^{(r)}}{\partial u_s} \frac{\partial u_s}{\partial x^{(t)}} &= 1 \quad \text{if } t = r \qquad\qquad (s \text{ umbral}) \\ &= 0 \quad \text{if } t \neq r \end{aligned}$$

we find that their product is unity and so we can write the quotient

$$\begin{aligned} \frac{d(y^{(1)} \cdots y^{(n)})}{d(x^{(1)} \cdots x^{(n)})} &\equiv \frac{\partial(y^{(1)} \cdots y^{(n)})}{\partial(u_1 \cdots u_n)} \cdot \frac{\partial(u_1 \cdots u_n)}{\partial(x^{(1)} \cdots x^{(n)})} \\ &\equiv \frac{\partial(y^{(1)} \cdots y^{(n)})}{\partial(x^{(1)} \cdots x^{(n)})} \text{ since } \frac{\partial y^{(r)}}{\partial u_m} \frac{\partial u_m}{\partial x^{(s)}} \equiv \frac{\partial y^{(r)}}{\partial x^{(s)}} \end{aligned}$$

$(m \text{ umbral})$

$$\equiv \left(\frac{\partial(x^{(1)} \cdots x^{(n)})}{\partial(y^{(1)} \cdots y^{(n)})}\right)^{-1} \text{ as above}$$
$$\equiv \sqrt{g/f} \text{ since } f = gJ^2.$$

Accordingly

$$\sqrt{f}\, d(y^{(1)} \cdots y^{(n)}) \equiv \sqrt{g}\, d(x^{(1)} \cdots x^{(n)})$$

so that this expression is an *invariant*. In view of the fact that it depends on the fundamental quadratic differential form $(ds)^2$ it is called a *metrical* invariant.

Let us consider an integral over a region of the fundamental space S_n, $\int X_{1 \dots n} d(x^{(1)} \cdots x^{(n)})$. Here $X_{1 \dots n}$ is the *single distinct* function of an *arbitrary* alternating covariant tensor of rank n. Since the integrand is invariant and since $\sqrt{g}\, d(x^{(1)} \cdots x^{(n)})$ is invariant it follows by division that $X_{1 \dots n} \div \sqrt{g}$ is an invariant. As an application of Stokes' Lemma we have already

$$(-1)^n X_{2 \dots n} \equiv X_{n2 \dots n-1}; \quad \cdots \quad (X_n) \equiv X_{1 \dots n-1}$$

(where $X_{s_1 \dots s_{n-1}}$ is any alternating covariant tensor of rank $n - 1$) then

$$X_{1 \dots n} \equiv \frac{\partial}{\partial x^{(s)}} (X_s) \qquad (s \text{ umbral})$$

is the coefficient of an integral over a region of S_n. We see therefore that $\dfrac{1}{\sqrt{g}} \dfrac{\partial}{\partial x^{(s)}} (X_s)$ is an *invariant.*

We shall now investigate the nature of the n functions (X_s). Under a transformation of coordinates from x to y we find, for example,

$$(Y_n) \equiv Y_{1 \dots n-1} \equiv X_{s_1 \dots s_{n-1}} \frac{\partial x^{(s_1)}}{\partial y^{(1)}} \cdots \frac{\partial x^{(s_{n-1})}}{\partial y^{(n-1)}}$$
$$(s_1 \cdots s_{n-1} \text{ umbral})$$
$$\equiv \sum_{s_1 < s_2 < \cdots < s_n} X_{s_1 \dots s_{n-1}} \frac{\partial(x^{(s_1)} \cdots x^{(s_{n-1})})}{\partial(y^{(1)} \cdots y^{(n-1)})}$$
$$(\text{owing to alternating character of } X_{s_1 \dots s_{n-1}})$$

And, accordingly, if we denote the cofactor of $\dfrac{\partial x^{(r)}}{\partial y^{(s)}}$ in the expansion of J by (J_{rs}) we have

$$(Y_n) \equiv (J_{sn})(X_s) \qquad (s \text{ umbral})$$

In general

$$(Y_r) \equiv (J_{sr})(X_s)$$

If we solve the n equations

$$\frac{\partial x^{(r)}}{\partial y^{(p)}} \frac{\partial y^{(p)}}{\partial x^{(s)}} \equiv \frac{\partial x^{(r)}}{\partial x^{(s)}} = 1 \quad \text{if } s = r \qquad (p \text{ umbral})$$
$$= 0 \quad \text{if } s \neq r \qquad r = 1 \cdots n$$

for $\dfrac{\partial y^{(p)}}{\partial x^{(s)}}$ we find

$$J \frac{\partial y^{(p)}}{\partial x^{(s)}} = (J_{sp})$$

so that we may write

$$(Y_r) = J(X_s) \frac{\partial y^{(r)}}{\partial x^{(s)}} \equiv \sqrt{\frac{f}{g}}\,(X_s) \frac{\partial y^{(r)}}{\partial x^{(s)}}$$

or

$$\frac{(Y_r)}{\sqrt{f}} \equiv \frac{(X_s)}{\sqrt{g}} \frac{\partial y^{(r)}}{\partial x^{(s)}} \qquad (s \text{ umbral})$$

showing that $\dfrac{(X_s)}{\sqrt{g}}$ is a contravariant tensor of rank one. We may then put $(X_s) = \sqrt{g}\, X^s$ and our previous result takes the form that $\dfrac{1}{\sqrt{g}} \dfrac{\partial}{\partial x^{(s)}} (\sqrt{g}\, X^s)$ is an invariant; X^s being any contravariant tensor of rank one. This metrical invariant is known as the *divergence* of the contravariant tensor.

5. Special results

If $u(x^{(1)} \cdots x^{(n)})$ is any invariant function of position the rule of differentiation

$$\frac{\partial u}{\partial y^{(r)}} \equiv \frac{\partial u}{\partial x^{(s)}} \frac{\partial x^{(s)}}{\partial y^{(r)}} \qquad (s \text{ umbral})$$

tells us that the n functions $X_s \equiv \dfrac{\partial u}{\partial x^{(s)}}$ form a covariant tensor of rank one; this is known as the *tensor gradient*. If X_r is any covariant tensor of rank one its simple product by itself or "square" is a covariant tensor of rank two, $X_{rs} \equiv X_r X_s$. Hence by rule (*d*), Ch. 2, § 4,

$$g^{rs}X_rX_s \text{ is an invariant} \qquad (r, s \text{ umbral})$$

This is called the square of the *magnitude* of the tensor. In particular the square of the tensor gradient is the invariant

$$\Delta_1 u \equiv g^{rs}\frac{\partial u}{\partial x^{(r)}}\frac{\partial u}{\partial x^{(s)}} \qquad (r, s \text{ umbral})$$

This is known as the "*first differential parameter of u.*" Similarly the magnitude of the square of a contravariant tensor of rank 1 is the invariant $g_{rs}X^{(r)}X^{(s)}$.

Again

$$g^{rs}\frac{\partial u}{\partial x^r} \equiv X^s \qquad (r \text{ umbral})$$

is contravariant of rank one (rule (*d*)). Hence

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial x^{(s)}}\left(\sqrt{g}\, g^{rs}\frac{\partial u}{\partial x^{(r)}}\right) \text{ is an invariant} \qquad (r, s \text{ umbral})$$

by the result of the preceding paragraph. It is written $\Delta_2 u$ and is known as the "second differential parameter."* In ordinary space of three dimensions in which the x's are rectangular Cartesian coordinates

$$\begin{aligned} g_{rs} &= 0 \qquad \text{if } r \neq s \\ &= 1 \qquad \text{if } r = s \end{aligned}$$

and $g_{rs} = g^{rs}$; $\sqrt{g} = 1$ so that $\Delta_2 u$ takes the form

$$\frac{\partial^2 u}{\partial x^{(1)2}} + \frac{\partial^2 u}{\partial x^{(2)2}} + \frac{\partial^2 u}{\partial x^{(3)2}}$$

* *Larmor*, *J.*, Transactions Cambridge Phil. Soc., Vol. 14, p. 121 (1885), obtains this transformation in the case $n = 3$ by the application of the Calculus of Variations.

When we change over to any "cu
we have under the form

$$\Delta_2 u \equiv \frac{1}{\sqrt{f}} \frac{\partial}{\partial y^{(s)}} \Big(\sqrt{f} f^{r}$$

the expression of this magnitude in a ro-
coordinates.

6. General orthogonal

Whenever we have, in any space, co
expression $(ds)^2$ involves only square te
the coordinates are said to be orthog
explained later). It is usual to write, in this case,

$$(ds)^2 \equiv \frac{1}{h_1^2}(dx_1)^2 + \frac{1}{h_2^2}(dx^{(2)})^2 + \cdots + \frac{1}{h_n^2}(dx^{(n)})^2;$$

accordingly

$$g_{11} = \frac{1}{h_1^2}; \qquad \cdots \qquad g_{nn} = \frac{1}{h_n^2}$$

so that

$$g = \frac{1}{h_1^2} \cdot \frac{1}{h_2^2} \cdots \frac{1}{h_n^2}; \qquad \sqrt{g} = \frac{1}{h_1 h_2 \cdots h_n}$$

$$g^{11} = h_1^2 \cdots g^{nn} = h_n^2; \qquad g^{rs} = 0 \quad r \neq s$$

The square of the gradient is

$$\Delta_1 u \equiv h_1^2 \left(\frac{\partial u}{\partial x^{(1)}} \right)^2 + \cdots + h_n^2 \left(\frac{\partial u}{\partial x^{(n)}} \right)^2$$

whilst the quantity

$$\Delta_2 u \equiv h_1 h_2 \cdots h_n \left\{ \frac{\partial}{\partial x^{(1)}} \left(\frac{h_1}{h_2 \cdots h_n} \frac{\partial u}{\partial x^{(1)}} \right) + \cdots + \frac{\partial}{\partial x^{(n)}} \left(\frac{h_n}{h_1 \cdots h_{n-1}} \frac{\partial u}{\partial x^{(n)}} \right) \right\}$$

The reader should write out the explicit formulæ for space polar and cylindrical coordinates in ordinary space of three dimensions.

7. The special or restricted vector analysis

In the form given to the theory by Heaviside and others only those coordinates x or y were considered in which the fundamental metrical form is

$$ds^2 = (dx_1^{(1)})^2 + \cdots + (dx^{(n)})^2 \equiv (dy^{(1)})^2 + \cdots + (dy^{(n)})^2$$

These coordinates we call rectangular or orthogonal Cartesian coordinates and the space we call Euclidean. It is true that use was made of Stokes' Lemma to find expressions for important invariants as $\Delta_2 u$ in other than orthogonal Cartesian coordinates but no attempt was made to define the components of a vector in these coordinates. Now when we restrict ourselves to that subgroup (of all the continuous transformations) which carries us from one set of orthogonal Cartesian coordinates to another the distinction between *covariant* and *contravariant* tensors completely disappears. The transformations are necessarily of the linear type

$$x^{(r)} = (a_{rs})y^{(s)} \quad (s \text{ umbral},\ r = 1 \cdots n)$$

where the a's are constants. Since here $f = g = 1$, $J^2 = 1$* and so the equations just written have a unique solution for the y's. To get this most conveniently note that $dx^{(r)} = (a_{rs})dy^{(s)}$ and squaring and adding we have

$$\begin{aligned} (a_{rs})(a_{rt}) &= 0 \qquad t \neq s \qquad\qquad (r \text{ umbral}) \\ &= 1 \qquad t = s \end{aligned}$$

Hence multiplying the equations for x by a_{rt} and using t as an umbral symbol we find

$$\begin{aligned} (a_{rt})x^{(r)} &\equiv (a_{rt})(a_{rs})y^{(s)} \qquad\qquad (r,\ s \text{ umbral}) \\ &\equiv y^{(t)} \end{aligned}$$

$$\therefore \quad \frac{\partial y^{(t)}}{\partial x^{(r)}} = (a_{rt}) = \frac{\partial x^{(r)}}{\partial y^{(t)}}\,.$$

Accordingly the equations of correspondence defining covariant

* We shall consider only *direct* transformations; those for which $J = +1$.

and contravariant tensors are, for this restricted set of transformations, identical. Again denoting by (A_{rs}) the cofactor of (a_{rs}) in the expansion of the determinant J we have by the usual method that

$$y^{(t)} = (A_{rt})x^{(r)}$$

and *since the solution is unique* we must have $(a_{rt}) = (A_{rt})$.* Hence since $g = 1$ we have that the n *distinct components* of an alternating tensor of rank $n - 1$ form *a tensor of rank one.* It is for this reason that when $n = 3$ it was found necessary to discuss but one kind of tensor—that of the first rank which was called a *vector*.† Still some writers felt a distinction between the two kinds; that of the first rank they called polar and the alternating tensor of the second kind, whose three distinct components form a tensor of the first kind, they called axial. Thus a velocity or gradient are polar vectors (the first being properly contravariant, the latter covariant) whilst a curl or a vector product are axial vectors.

When, in the mathematical discussion of the *Special Relativity Theory,* it was found convenient to make $n = 4$ [the transformations (Lorentz) being still those of the linear orthogonal type], a new kind of tensor or vector is introduced. Here it is the alternating tensor of the third rank which, when we consider merely its four distinct components, is equivalent, from its definition and the properties of the transformation, to a tensor of the first rank or " four-vector." There remains the alternating tensor of the second rank and the six distinct components of this were known, for want of a better name, as a six-vector. As an example of the general theory we have that

(a) the divergence of a four-vector $\dfrac{\partial X^{(s)}}{\partial x^{(s)}}$ is an invariant.

(s umbral)

* This is merely a special case of the previous result that $J\dfrac{\partial y^{(r)}}{\partial x^{(s)}} = (J_{sr})$.

† Until a consideration of non-alternating tensors became desirable.

tells us that the n functions $X_s \equiv \dfrac{\partial u}{\partial x^{(s)}}$ form a covariant tensor of rank one; this is known as the *tensor gradient*. If X_r is any covariant tensor of rank one its simple product by itself or "square" is a covariant tensor of rank two, $X_{rs} \equiv X_r X_s$. Hence by rule (d), Ch. 2, § 4,

$$g^{rs} X_r X_s \text{ is an invariant} \qquad (r, s \text{ umbral})$$

This is called the square of the *magnitude* of the tensor. In particular the square of the tensor gradient is the invariant

$$\Delta_1 u \equiv g^{rs} \frac{\partial u}{\partial x^{(r)}} \frac{\partial u}{\partial x^{(s)}} \qquad (r, s \text{ umbral})$$

This is known as the "*first differential parameter of u.*" Similarly the magnitude of the square of a contravariant tensor of rank 1 is the invariant $g_{rs} X^{(r)} X^{(s)}$.

Again

$$g^{rs} \frac{\partial u}{\partial x^r} \equiv X^s \qquad (r \text{ umbral})$$

is contravariant of rank one (rule (d)). Hence

$$\frac{1}{\sqrt{g}} \frac{\partial}{\partial x^{(s)}} \left(\sqrt{g}\, g^{rs} \frac{\partial u}{\partial x^{(r)}} \right) \text{ is an invariant} \qquad (r, s \text{ umbral})$$

by the result of the preceding paragraph. It is written $\Delta_2 u$ and is known as the "second differential parameter."* In ordinary space of three dimensions in which the x's are rectangular Cartesian coordinates

$$\begin{aligned} g_{rs} &= 0 \qquad \text{if } r \neq s \\ &= 1 \qquad \text{if } r = s \end{aligned}$$

and $g_{rs} = g^{rs}$; $\sqrt{g} = 1$ so that $\Delta_2 u$ takes the form

$$\frac{\partial^2 u}{\partial x^{(1)2}} + \frac{\partial^2 u}{\partial x^{(2)2}} + \frac{\partial^2 u}{\partial x^{(3)2}}$$

* *Larmor*, *J.*, Transactions Cambridge Phil. Soc., Vol. 14, p. 121 (1885), obtains this transformation in the case $n = 3$ by the application of the Calculus of Variations.

When we change over to any "curvilinear" coordinates y we have under the form

$$\Delta_2 u \equiv \frac{1}{\sqrt{f}} \frac{\partial}{\partial y^{(s)}} \left(\sqrt{f} f^{rs} \frac{\partial u}{\partial y^{(s)}} \right)$$

the expression of this magnitude in a form suited to the new coordinates.

6. General orthogonal coordinates

Whenever we have, in any space, coordinates x such that the expression $(ds)^2$ involves only square terms, i.e., $g_{rs} \equiv 0$ if $s \neq r$, the coordinates are said to be orthogonal (for a reason to be explained later). It is usual to write, in this case,

$$(ds)^2 \equiv \frac{1}{h_1^2} (dx_1)^2 + \frac{1}{h_2^2} (dx^{(2)})^2 + \cdots + \frac{1}{h_n^2} (dx^{(n)})^2;$$

accordingly

$$g_{11} = \frac{1}{h_1^2}; \qquad \cdots \qquad g_{nn} = \frac{1}{h_n^2}$$

so that

$$g = \frac{1}{h_1^2} \cdot \frac{1}{h_2^2} \cdots \frac{1}{h_n^2}; \qquad \sqrt{g} = \frac{1}{h_1 h_2 \cdots h_n}$$

$$g^{11} = h_1^2 \cdots g^{nn} = h_n^2; \qquad g^{rs} = 0 \quad r \neq s$$

The square of the gradient is

$$\Delta_1 u \equiv h_1^2 \left(\frac{\partial u}{\partial x^{(1)}} \right)^2 + \cdots + h_n^2 \left(\frac{\partial u}{\partial x^{(n)}} \right)^2$$

whilst the quantity

$$\Delta_2 u \equiv h_1 h_2 \cdots h_n \left\{ \frac{\partial}{\partial x^{(1)}} \left(\frac{h_1}{h_2 \cdots h_n} \frac{\partial u}{\partial x^{(1)}} \right) + \cdots + \frac{\partial}{\partial x^{(n)}} \left(\frac{h_n}{h_1 \cdots h_{n-1}} \frac{\partial u}{\partial x^{(n)}} \right) \right\}$$

The reader should write out the explicit formulæ for space polar and cylindrical coordinates in ordinary space of three dimensions.

(b) From any six-vector X_{rs} we may derive a four-vector (really an alternating tensor of the third rank)

$$X_{rst} \equiv \frac{\partial X_{rs}}{\partial x^{(t)}} + \frac{\partial X_{st}}{\partial x^{(r)}} + \frac{\partial X_{tr}}{\partial x^{(s)}}$$

It is this four-vector that was written lor X_{rs} in honor of Lorentz.

8. Generalization of the reciprocal relationship between an alternating tensor of rank r and one of opposite kind of rank $n - r$ from the case $r = 1$ already treated to a general value of r.

We have already seen that

$$J\frac{\partial y^{(p)}}{\partial x^{(s)}} = (J_{sp})$$

where J is the determinant $\dfrac{\partial(x^{(1)} \cdots x^{(n)})}{\partial(y^{(1)} \cdots y^{(n)})}$ of the transformation ...d (J_{sp}) is the cofactor, in the expansion of J, of the element $(j_{sp}) \equiv \dfrac{\partial x^{(s)}}{\partial y^{(p)}}$ of this determinant.

Hence

$$J^2 \begin{vmatrix} \dfrac{\partial y^{(r_1)}}{\partial x^{(s_1)}} & \dfrac{\partial y^{(r_1)}}{\partial x^{(s_2)}} \\ \dfrac{\partial y^{(r_2)}}{\partial x^{(s_1)}} & \dfrac{\partial y^{(r_2)}}{\partial x^{(s_2)}} \end{vmatrix} \equiv \begin{vmatrix} J_{s_1r_1} & J_{s_2r_1} \\ J_{s_1r_2} & J_{s_2r_2} \end{vmatrix}$$

Now the determinant of the minors of J is well known to be equivalent to the product of J by the determinant of order $n - 2$ obtained by erasing the s_1th and s_2th rows and the r_1th and r_2th columns of J affected with its proper sign (the determinant of order $n - 2$ is the cofactor of $\begin{vmatrix} \dfrac{\partial x^{(s_1)}}{\partial y^{(r_1)}} & \dfrac{\partial x^{(s_2)}}{\partial y^{(r_1)}} \\ \dfrac{\partial x^{(s_1)}}{\partial y^{(r_2)}} & \dfrac{\partial x^{(s_2)}}{\partial y^{(r_2)}} \end{vmatrix}$ in the Laplacian ex-

pansion of J in terms of two row determinants from the s_1th and s_2th rows and the r_1th and r_2th rows). Hence we have the result that the $n(n-1)/2$ distinct components of an alternating covariant tensor of rank $n-2$ when *divided* by $\sqrt{g}$ form the distinct components of an alternating contravariant tensor of rank two. And so in general. Similarly the $\binom{n}{r}$ distinct components of an alternating contravariant tensor of rank $n-r$ when *multiplied* by $\sqrt{g}$ form an alternating covariant tensor of rank r.

Example. Take $n = 4$, $r = 2$ and consider the linear orthogonal transformations of the Special Relativity Theory. Here

$$\begin{array}{lll} X_{12} = X^{34}; & X_{13} = X^{42}; & X_{14} = X^{23} \\ X_{23} = X^{14}; & X_{24} = X^{31}; & X_{34} = X^{12} \end{array}$$

The two tensors or six vectors X_{rs} and X^{rs} were said to be *reciprocal.**

* Cf. *Cunningham, E., The Principle of Relativity*, Ch. 8, Camb. Univ. Press (1914).

CHAPTER IV

1. Geometrical interpretation of the components of a tensor

Definitions

(a) Direction of a curve at any point on it.

At any point u on the *curve* V_1 specified by the equations

$$x^{(s)} \equiv x^{(s)}(u) \qquad (s = 1, \cdots, n)$$

whose length s from a fixed point u_0 is defined by the integral

$$s \equiv \int_{u_0}^{u} \sqrt{g_{lm} \frac{dx^{(l)}}{du} \frac{dx^{(m)}}{du}}\, du$$

we may form the n quantities

$$l^{(r)} \equiv \frac{dx^{(r)}}{ds} \equiv \frac{dx^{(r)}}{du} \div \frac{ds}{du} \qquad (r = 1, \cdots, n)$$

We exclude from consideration here the "minimal" curves along which $ds = 0$. Since $X^r \equiv dx^{(r)}$ is a contravariant tensor of rank one and ds is an invariant we have that the n quantities $l^{(r)}$ form a contravariant tensor of rank one which we call the "*direction*" tensor of the curve at the point in question. The n components we call *direction coefficients.* The equation of definition

$$(ds)^2 \equiv g_{rs} dx^{(r)} dx^{(s)} \qquad (r, s \text{ umbral})$$

shows us that $g_{rs} l^{(r)} l^{(s)} = 1$ so that a knowledge of the mutual ratios of the direction coefficients suffices to determine their magnitudes (save for an indefiniteness as to sign). Otherwise expressed the *magnitude of the direction tensor is unity.* Fixing the indefiniteness as to sign by a particular choice is said to fix

a *sense of direction on the curve* and the curve may be then said to be directed.

2. (*b*) *Metrical Definition of Angle*

Consider two curves with a common point and let their direction tensors at this point be $l^{(r)}$ and $m^{(r)}$. The simple product $X^{rs} \equiv l^{(r)}m^{(s)}$ is contravariant of rank two (Rule (*c*), Ch. 2) and so the expression $g_{rs}l^{(r)}m^{(s)}$ is invariant (r, s umbral; Rule (*d*), Ch. 2). This we call the cosine of the angle θ between the two curves (directed) at the point. If the quadratic differential form defining $(ds)^2$ is supposed to be *definite*, i.e., if it is supposed that (ds) cannot be zero, for *real* values of the variables $x^{(r)}$ and $dx^{(r)}$ save in the trivial case when all the $dx^{(r)} = 0$, it can easily be shown that the angle defined in this way is always real for real curves. Let us write instead of $dx^{(r)}$ the expression $\lambda l^{(r)} + \mu m^{(r)}$ and thus form the quadratic expression in λ and μ

$$g_{rs}(\lambda l^{(r)} + \mu m^{(r)})(\lambda l^{(s)} + \mu m^{(s)})$$

This is not to vanish for *real* values of λ, μ save when $\lambda = 0$, $\mu = 0$ (we suppose the quantities $l^{(r)}$ and $m^{(r)}$ all real and the two directions as distinct). Using

$$g_{rs}l^{(r)}l^{(s)} = 1 = g_{rs}m^{(r)}m^{(s)}$$

we have that

$$\lambda^2 + 2\lambda\mu \cos\theta + \mu^2 = 0$$

must have complex roots when regarded as an equation in $\lambda : \mu$. Hence $1 - \cos^2\theta > 0$ so that the angle as defined above is always *real* for *real* directions under the assumption that (ds) cannot vanish on a *real* curve. It must be remembered however that this assumption is not always made, e.g., in Relativity Theory.

When $\cos\theta = 0$ the curves are said to be *orthogonal* or *at right angles* at the point in question.

EXAMPLES

In ordinary space with the x's as rectangular Cartesian coordinates we have the usual expression

$$\cos\theta = l^{(1)}m^{(1)} + l^{(2)}m^{(2)} + l^{(3)}m^{(3)}$$

where $(l^{(1)}, l^{(2)}, l^{(3)})$, $(m^{(1)}, m^{(2)}, m^{(3)})$ are the direction *cosines* of the two curves. If now we use any "curvilinear" coordinates $(y^{(1)}, y^{(2)}, y^{(3)})$ the angle between two curves is

$$\cos\theta = f_{rs}\left(\frac{dy^{(r)}}{ds}\right)_1\left(\frac{dy^{(s)}}{ds}\right)_2 \qquad (r, s \text{ umbral})$$

In particular if we use orthogonal coordinates

$$(ds)^2 = f_{11}(dy^{(1)})^2 + f_{22}(dy^{(2)})^2 + f_{33}(dy^{(3)})^2$$

$$\cos\theta = f_{11}\left(\frac{dy^{(1)}}{ds}\right)_1\left(\frac{dy^{(1)}}{ds}\right)_2 + \text{etc.}$$

Thus for a curve in polar coordinates r, θ, ϕ

$$\cos\theta = \left(\frac{dr}{ds}\right)_1\left(\frac{dr}{ds}\right)_2 + r^2\left(\frac{d\theta}{ds}\right)_1\left(\frac{d\theta}{ds}\right)_2 + r^2\sin^2\theta\left(\frac{d\phi}{ds}\right)_1\left(\frac{d\varphi}{ds}\right)_2$$

It will now be clear why those coordinates in terms of which $(ds)^2$ has no product terms are said to be *orthogonal*.

For

$$f_{rs} \equiv g_{lm}\frac{\partial x^l}{\partial y^{(r)}}\frac{\partial x^{(m)}}{\partial y^{(s)}} \qquad \text{(from its covariant character)}$$

If now all the coordinates y but one, $y^{(r)}$ say, are kept constant we have a curve whose equations, in the x coordinates, may be conveniently specified by means of the parameter $y^{(r)}$

$$x^{(s)} \equiv x^{(s)}(y^{(r)}) \qquad (s = 1, \cdots, n)$$

Through each point y there pass n curves of this kind which we shall call the n *coordinate lines* y through that point. On the rth of these coordinate lines the direction tensor is

$$X^l = \frac{dx^{(l)}}{ds} = \frac{\partial x^{(l)}}{\partial y^{(r)}} \div \frac{ds}{dy^{(r)}}$$

and so the vanishing of the component f_{rs} states that the coordinate *lines* $y^{(r)}$ and $y^{(s)}$ are orthogonal. Hence if $(ds)^2$ does not contain any product terms the coordinate lines are *everywhere*, all *mutually orthogonal* and so the coordinates are said to be orthogonal. In ordinary space, i.e., where the x's are rectangular Cartesian coordinates and where the y's are orthogonal coordinates,

$$f_{11} = \Sigma_r \left(\frac{\partial x^{(r)}}{\partial y^{(1)}} \right)^2$$

$$f^{11} = \Sigma_r \left(\frac{\partial y^{(1)}}{\partial x^{(r)}} \right)^2$$

and

$$f_{11} = (f^{11})^{-1}$$

so that

$$\Sigma_r \left(\frac{\partial x^{(r)}}{\partial y^{(1)}} \right)_2 = 1 \div \Sigma_r \left(\frac{\partial y^{(1)}}{\partial x^{(r)}} \right)^2$$

a result which is sometimes useful in the calculation of the coefficients f_{11}, f_{22}, $f_{33} \cdots$ of the form $(ds)^2$ in the curvilinear coordinates y.

3. Resolution of tensors

If we consider any *covariant* tensor X_r of rank one and take the inner product of this into a direction tensor $l^{(r)}$ we derive the invariant $X_r l^{(r)}$ (r umbral; Rule (d)). This we call the *resolved part* of the covariant tensor along the direction $l^{(r)}$. Let us now make a transformation of coordinates from x to y and consider the coordinate line $y^{(s)}$. The n components of the direction tensor for this curve are proportional to

$$\frac{\partial x^{(r)}}{\partial y^{(s)}} \qquad (r = 1, \cdots, n)$$

To determine the actual values of these components we must divide through by the positive square root of

$$g_{lm} \frac{\partial x^{(l)}}{\partial y^{(s)}} \frac{\partial x^{(m)}}{\partial y^{(s)}} \qquad (l, m \text{ umbral})$$

and this is equivalent to $\sqrt{f_{ss}}$.

The equations defining the covariant correspondence for a tensor of the first rank are

$$Y_l \equiv X_r \frac{\partial x^{(r)}}{\partial y^{(l)}} \qquad (l = 1, \cdots, n;\ r \text{ umbral})$$

$= \sqrt{f_{ll}}$ times the resolved part of the tensor X_r along the coordinate direction $y^{(l)}$

Example

Space polar coordinates y in ordinary space of three dimensions. The x are rectangular Cartesian coordinates. Denoting the resolved parts of the *covariant* tensor X in the directions $y^{(1)}, y^{(2)}, y^{(3)}$ by R, Θ, Φ respectively we have since $f_{11} = 1$; $f_{22} = r^2$; $f_{33} = r^2 \sin^2 \theta$

$$Y_1 \equiv R; \qquad Y_2 \equiv r\Theta; \qquad Y_3 \equiv r \sin \theta \Phi.$$

ree distinct components of the alternating covariant tensor k two, curl X, in polar coordinates are

$$\frac{\partial}{\partial \theta}(r \sin \theta \Phi) - \frac{\partial}{\partial \phi}(r\Theta)$$
$$\frac{\partial R}{\partial \phi} - \frac{\partial}{\partial r}(r \sin \theta \Phi)$$
$$\frac{\partial}{\partial r}(r\Theta) - \frac{\partial R}{\partial \theta}$$

Similarly for cylindrical coordinates ρ, ϕ, z where $f_{11} = 1$; $f_{22} = \rho^2$; $f_{33} = 1$ if we denote the resolved parts of X along the three coordinate directions by R, Φ, Z we have $Y_1 \equiv R$; $Y_2 \equiv \rho\Phi$; $Y_3 \equiv Z$ and the components of the curl are at once written down.

Resolution of Contravariant Tensors.

To define what is meant by this we require, not as before the coordinate lines $y^{(r)}$ along each of which all the coordinates y but one, $y^{(r)}$, are constant, but the coordinate *spreads* V_{n-1} along each

of which all the variables but one, $y^{(r)}$ say, vary. The parameters $u_1 \cdots u_{n-1}$ may here be very conveniently chosen to be the coordinates $y_1 \cdots y_n$ themselves omitting $y^{(r)}$, and then $y^{(r)}$ is a constant (a particular function of $u_1 \cdots u_{n-1}$). Now, in general, when we have a V_{n-1} specified by equations

$$x^{(s)} \equiv x^{(s)}(u_1, \cdots, u_{n-1}) \qquad (s = 1, \cdots, n)$$

we obtain on the spread, through each point, $n - 1$ parameter lines by letting in turn each parameter vary, keeping all the rest fixed. Any one of these, u_r varying, say, has at the point in question a direction tensor whose components are proportional to

$$\frac{\partial x^{(s)}}{\partial u_r} \qquad (s = 1, \cdots, n)$$

Let us look for a direction orthogonal at once to the $n - 1$ directions of these parameter curves. Such a direction tensor has components $n^{(1)} \cdots n^{(n)}$ say and is said to be normal to the spread V_{n-1} at the point in question. To express the required orthogonality we have $n - 1$ equations

$$g_{lm}n^{(l)} \frac{\partial x^{(m)}}{\partial u_r} = 0 \quad (l,\, m \text{ umbral};\ \ r = 1 \cdots n)$$

homogeneous in the $n^{(1)} \cdots n^{(n)}$ and thus serving to determine their mutual ratios. To actually solve divide across by one of the unknowns $n^{(n)}$ say and we have $n - 1$ linear, non-homogeneous equations for the $(n - 1)$ unknowns

$$v_1 = \frac{n^{(1)}}{n^{(n)}}, \qquad \cdots, \qquad v_{n-1} = \frac{n^{(n-1)}}{n^{(n)}} \text{ *}$$

* The algebra following here is somewhat complicated and so it may be desirable to derive the expressions for the components of the normal direction tensor to the spread $y^{(n)}$ as follows. Working *with the coordinates* y the $n - 1$ parameter curves $y^{(s)}$ varying ($s = 1, \cdots, n - 1$) have their direction coefficients proportional to

$$\left.\begin{matrix}(1, & 0 & & & 0)\\ (0, & 1, & 0 & & 0)\end{matrix}\right\}$$

$$\begin{matrix}(0, & 0 & & 1, & 0)\end{matrix}$$

The determinant of the coefficients has as the element in the rth row and sth column

$$g_{sm}\frac{\partial x^{(m)}}{\partial u_r} \qquad (m \text{ umbral; } r, s = 1, \cdots, n-1)$$

This determinant is therefore the product of the two matrices

$$\left\|\begin{matrix} g_{11} & g_{12} & \cdots & g_{1n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ g_{n-1,1} & & \cdots & g_{n-1,n} \end{matrix}\right\| \left\|\begin{matrix} \dfrac{\partial x^{(1)}}{\partial u_1} & \dfrac{\partial x^{(2)}}{\partial u_1} & \cdots & \dfrac{\partial x^{(n)}}{\partial u_1} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \dfrac{\partial x^{(1)}}{\partial u_{n-1}} & & \cdots & \dfrac{\partial x^{(n)}}{\partial u_{n-1}} \end{matrix}\right\|$$

each of $n-1$ rows and n columns. It is well known that this product can also be written as the sum of products of all corre-

respectively. The $n-1$ equations expressing that $n^{(r)}$ is orthogonal to these $n-1$ directions are

$$f_{tr}n^{(r)} = 0 \qquad (t = 1, \cdots, n-1;\ r \text{ umbral})$$

Hence the ratios

$$n^{(1)} : n^{(2)} : \cdots : n^{(n)} = f^{n1} : f^{n2} : \cdots : f^{nn}$$

the actual values being these divided by

$$\sqrt{f_{rs}f^{nr}f^{ns}} = \sqrt{f^{nn}}$$

[one must be warned against thinking that $\dfrac{f^{ns}}{\sqrt{f^{nn}}}$ $(s = 1, \cdots, n)$ are contravariant. When a change of coordinates from y to x is made the spread $y^{(n)} = \text{const.}$ does not become $x^{(n)} = \text{const.}$] If now we wish to use x coordinates, the normal direction tensor, being contravariant of rank one, has components proportional to

$$\begin{aligned} n^{(r)} &= f^{ns}\frac{\partial x^{(r)}}{\partial y^{(s)}} && (r = 1, \cdots, n;\ s \text{ umbral}) \\ &= g^{lm}\frac{\partial y^n}{\partial x^l}\frac{\partial y^{(s)}}{\partial x^m}\frac{\partial x^r}{\partial y^s} && (l, m, s \text{ umbral}) \\ &= g^{lr}\frac{\partial y^{(n)}}{\partial x^{(l)}} && (l \text{ umbral}) \end{aligned}$$

If $y^{(n)} = V(x^{(1)}, \cdots, x^{(n)})$ *we have that the normal direction tensor to the spread has its components proportional to* $g^{lr}\dfrac{\partial V}{\partial x^{(l)}}$ the result required.

sponding determinants of order $n - 1$ that can be formed from each matrix. Let us write for brevity

$$(J_1) \equiv (-1)^{n-1}\frac{\partial(x^{(2)} \cdots x^{(n)})}{\partial(u_1 \cdots u_{n-1})}$$
$$(J_2) \equiv (-1)^{n-2}\frac{\partial(x^{(1)}\, x^{(3)} \cdots x^{(n)})}{\partial(u_1 \cdots u_{n-1})}$$
$$(J_n) \equiv \frac{\partial(x^{(1)}\, x^{(2)} \cdots x^{(n-1)})}{\partial(u_1 \cdots u_{n-1})}$$

and the determinant of the coefficients becomes

$$(G_{ns})(J_s) \qquad (s \text{ umbral})$$

which may be written $g\, g^{ns}(J_s)$. The numerators of the fractions furnishing $v_1 \cdots v_{n-1}$ are dealt with in the same way and we have

$$n^{(1)} : n^{(2)} : \cdots : n^{(n)} \equiv g^{1s}(J_s) : g^{2s}(J_s) : \cdots : g^{ns}(J_s)$$

(Since the (J_s) are really the n distinct components of an *alternating contravariant* tensor of rank $n - 1$ we know that $X_s \equiv (J_s)\sqrt{g}$ is a covariant tensor of rank one verifying the contravariant character of the $n^{(r)}$ (Rule (*d*))). If all the (J_s) vanish the point is said to be a singular point of the spread and the determination of $n^{(r)}$ becomes impossible.

Let us now apply these generalities to the spread V_{n-1} given by a single equation

$$V(x^{(1)} \cdots x^{(n)}) = 0$$

connecting the coordinates x. We may solve for one of the coordinates, $x^{(n)}$ say, in terms of the others $x^{(1)} \cdots x^{(n-1)}$ and these others we use as the $n - 1$ independent parameters of the spread:

$$x^{(1)} \equiv u_1 \cdots x^{(n-1)} \equiv u_{n-1} \qquad x^{(n)} \equiv x^{(n)}(u_1 \cdots u_{n-1})$$

are then the equations, in parametric form, of the spread V_{n-1}. Our matrix

$$\frac{\partial x^{(r)}}{\partial u_s}\ (r = 1, \cdots, n; s = 1, \cdots, n - 1)$$

is now

$$\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & \dfrac{\partial x^{(n)}}{\partial x^{(1)}} \\ 0 & 1 & & \cdots & & \dfrac{\partial x^{(n)}}{\partial x^{(2)}} \\ & & & & & \\ 0 & 0 & & \cdots & 1 & \dfrac{\partial x^{(n)}}{\partial x^{(n-1)}} \end{vmatrix}$$

and so

$$(J_1) \equiv (-1)^{n-1} \frac{\partial(x^{(2)} \cdots x^{(n)})}{\partial(u_1, \cdots, u_{n-1})} \equiv - \frac{\partial(x^{(n)}\ x^{(2)} \cdots x^{(n-1)})}{\partial(x^{(1)} \cdots x^{(n-1)})}$$

$$\equiv - \begin{vmatrix} \dfrac{\partial x^{(n)}}{\partial x^{(1)}} & 0 & \cdots & 0 \\ \cdot & 1 & \cdots & 0 \\ \cdot & 0 & \cdots & \cdot \\ & & & \\ \dfrac{\partial x^{(n)}}{\partial x^{(n-1)}} & & \cdots & 1 \end{vmatrix} \equiv - \frac{\partial x^{(n)}}{\partial x^{(1)}}$$

But, on differentiating the equation $V(x^{(1)} \cdots x^{(n)}) = 0$ of our spread V_{n-1} we obtain

$$\frac{\partial V}{\partial x^{(1)}} + \frac{\partial V}{\partial x^{(n)}} \frac{\partial x^{(n)}}{\partial x^{(1)}} = 0$$

so that

$$\frac{\partial x^{(n)}}{\partial x^{(1)}} = - \frac{\partial V}{\partial x^{(1)}} \div \frac{\partial V}{\partial x^{(n)}}$$

whence

$$(J_1) : (J_2) : \cdots : (J_n) = \frac{\partial V}{\partial x^{(1)}} : \frac{\partial V}{\partial x^{(2)}} : \cdots : \frac{\partial V}{\partial x^{(n)}}$$

In particular, if the spread V_{n-1} has, in the y coordinates, the equation $y^{(r)} = \text{const.}$, we have for its normal direction tensor

$$n^{(1)} : n^{(2)} : \cdots : n^{(n)} = g^{1s} \frac{\partial y^{(r)}}{\partial x^{(s)}} : \cdots : g^{ns} \frac{\partial y^{(r)}}{\partial x^{(s)}} \quad (s \text{ umbral})$$

The actual magnitudes of these components are found by dividing

through by the positive square root of

$$g_{lm} g^{ls} \frac{\partial y^{(r)}}{\partial x^{(s)}} g^{mt} \frac{\partial y^{(r)}}{\partial x^{(t)}}. \qquad (r \text{ not umbral})$$

$$\text{which expression is} \equiv g^{mt} \frac{\partial y^{(r)}}{\partial x^{(m)}} \frac{\partial y^{(r)}}{\partial x^{(t)}} \qquad (r \text{ not umbral})$$

$$\equiv f^{rr}$$

If now we have a contravariant tensor $X^{(r)}$ of rank one it is meaningless to call $X^{(r)} l^{(r)}$ the resolved part of the tensor in the direction l for the simple reason that this expression is not *invariant* but takes on different values in different systems of coordinates. However, we may first form the covariant tensor

$$X_s \equiv g_{sr} X^{(r)} \qquad (r \text{ umbral. Rule } (d))$$

This tensor is said to be *reciprocal* to the contravariant tensor $X^{(r)}$ with respect to the fundamental metrical quadratic differential form and *its* resolved part in any direction we call the *resolved part* of the *contravariant tensor in that direction*. Thus, for example, the resolved part of the contravariant tensor X^r in the direction normal to the coordinate spread $y^{(r)}$ = constant is

$$g_{sp} X^p \cdot n^{(s)} \equiv g_{sp} X^p g^{st} \frac{\partial y^{(r)}}{\partial x^{(t)}} \div \sqrt{f^{rr}} \qquad (s, p, t \text{ umbral})$$

$$\equiv \frac{1}{\sqrt{f^{rr}}} X^{(t)} \frac{\partial y^{(r)}}{\partial x^{(t)}} \qquad (t \text{ umbral})$$

$$\equiv \frac{1}{\sqrt{f^{rr}}} Y^{(r)}$$

Hence any component $Y^{(r)}$ of a contravariant tensor of rank one is the product by $\sqrt{f^{rr}}$ of the resolved part of the contravariant tensor normal to the coordinate spread $y^{(r)}$ = constant. It is now apparent that to deal with covariant and contravariant tensors of the first rank we require the coordinate lines through each point and the normals to the coordinate spreads through that point. *When the coordinates are orthogonal,* and only then,

these lines and normals coincide and a great simplification is due to this fact. This explains why orthogonal coordinates have been used, almost to the point of excluding all others, in the investigations of Theoretical Physics.

4. Example (a)

Space polar coordinates. These being orthogonal the normals to the spreads $r =$ const., $\theta =$ const., $\phi =$ constant are the coordinate lines r, θ, ϕ respectively and, if we denote the resolved parts of the contravariant tensor $X^{(s)}$ in these directions by R, Θ, Φ the three components are

$$Y^{(1)} = R; \qquad Y^{(2)} = \frac{\Theta}{r}; \qquad Y^{(3)} = \frac{\Phi}{r \sin \theta}.$$

In general for orthogonal coordinates y with

$$(ds)^2 = f_{11}(dy^{(1)})^2 + \cdots + f_{nn}(dy^{(n)})^2$$

we have $f^{rr} = 1/f_{rr}$ and if, as usual, we write $f_{rr} \equiv 1/h_r^2$ we have

$$f = (h_1^2 h_2^2 \cdots h_n^2)^{-1} \qquad \text{and} \qquad f^{rr} \equiv h_r^2$$

Here $Y^{(1)} \equiv h_1(R_1) \cdots Y^{(n)} = h_n(R_n)$ where we denote by $(R_1) \cdots (R_n)$ the resolved parts of the contravariant tensor along the coordinate directions $1, 2, \cdots, n$ respectively. The *divergence* of the contravariant tensor

$$\frac{1}{\sqrt{g}} \frac{\partial}{\partial x^{(s)}} \{\sqrt{g} X^s\}$$

takes the form

$$h_1 \cdots h_n \left\{ \frac{\partial}{\partial y^{(1)}} \left(\frac{(R_1)}{h_2 \cdots h_n} \right) + \cdots + \frac{\partial}{\partial y^{(n)}} \left(\frac{(R_n)}{h_1 \cdots h_{n-1}} \right) \right\}$$

Thus, for space polar coordinates, the divergence is

$$\frac{1}{r^2 \sin \theta} \left\{ \frac{\partial}{\partial r} (r^2 \sin \theta R) + \frac{\partial}{\partial \theta} (r \sin \theta \Theta) + \frac{\partial}{\partial \phi} (r\Phi) \right\}$$

and for cylindrical

$$\frac{1}{\rho}\left\{\frac{\partial}{\partial\rho}(\rho R)+\frac{\partial\Phi}{\partial\phi}+\frac{\partial}{\partial z}(\rho Z)\right\}$$

Example (*b*)

In order to illustrate the distinction between covariant and contravariant tensors we now consider *oblique* Cartesian coordinates y so that

$$(ds)^2 \equiv (dy^{(1)})^2 + (dy^{(2)})^2 + (dy^{(3)})^2 + 2\lambda dy^{(2)}dy^{(3)} + 2\mu dy^{(3)}dy^{(1)} + 2\nu dy^{(1)}dy^{(2)}$$

where the *constants* λ, μ, ν are the cosines of the angles between the oblique directed axes. Here

$$f \equiv \begin{vmatrix} 1 & \nu & \mu \\ \nu & 1 & \lambda \\ \mu & \lambda & 1 \end{vmatrix} = \text{square of volume of unit parallelopiped with its edges along the three axes.}$$

i.e., $\sqrt{f} = \sin\lambda\cos\theta_1 = \sin\mu\cos\theta_2 = \sin\nu\cos\theta_3$ where θ_1 is the angle between the coordinate line y_1 and the normal n_1 to the coordinate plane $y_1 =$ const. with similar definitions for θ_2 and θ_3.

Hence

$$\sqrt{f^{11}} \equiv \sqrt{\frac{1-\lambda^2}{f}} = \sec\theta_1; \qquad \sqrt{f^{22}} = \sec\theta_2; \qquad \sqrt{f^{33}} = \sec\theta_3$$

If we have any vector whose components in *rectangular* Cartesian coordinates $(x^{(1)}, x^{(2)}, x^{(3)})$ are X_1, X_2, X_3 this vector may be regarded as either a covariant or contravariant tensor, i.e., $X_1 \equiv X^{(1)}$; $X_2 \equiv X^{(2)}$; $X_3 \equiv X^{(3)}$ and if we denote the resolved parts of this vector along the coordinate *lines* y by $(X_{l_1}, X_{l_2}, X_{l_3})$ and along the normals of the coordinate *planes* y by $(X_{n_1}, X_{n_2}, X_{n_3})$ we have

$$Y_1 \equiv \sqrt{f_{11}}\,X_{l_1} \equiv X_{l_1}; \qquad Y_2 \equiv X_{l_2}; \qquad Y_3 \equiv X_{l_3}$$
$$Y^{(1)} \equiv \sqrt{f^{11}}\,X_{n_1} \equiv X_{n_1}\sec\theta_1; \qquad Y^{(2)} \equiv X_{n_2}\sec\theta_2; \qquad Y^3 \equiv X_{n_3}\sec\theta_3$$

Hence (Y_1, Y_2, Y_3) are the *resolved parts* of the vector along the three coordinate lines whilst $(Y^{(1)}, Y^{(2)}, Y^{(3)})$ are the *components* of the vector along these same directions. The tensors Y_r and Y^r are reciprocal with respect to the differential form $(ds)^2$, i.e.,

$$Y_1 = Y^{(1)} + \nu Y^{(2)} + \mu Y^{(3)}, \qquad \text{etc.}$$

Let us now consider the contravariant tensor whose components are

$$\overline{Y}^{(1)} \equiv \rho_1 Y^{(1)}; \qquad \overline{Y}^{(2)} \equiv \rho_2 Y^{(2)}; \qquad \overline{Y}^{(3)} \equiv \rho_3 Y^{(3)}$$

where ρ_1, ρ_2, ρ_3 are scalar or invariant numbers; we find for the components in the rectangular coordinate system x

$$\begin{aligned}\overline{X}^{(1)} &\equiv \overline{Y}^{(1)} \frac{\partial x^{(1)}}{\partial y^{(1)}} + \overline{Y}^{(2)} \frac{\partial x^{(1)}}{\partial y^{(2)}} + \overline{Y}^{(3)} \frac{\partial x^{(1)}}{\partial y^{(3)}} \\ &\equiv \rho_1 \left(X^{(1)} \frac{\partial y^{(1)}}{\partial x^{(1)}} + X^{(2)} \frac{\partial y^{(1)}}{\partial x^{(2)}} + X^{(3)} \frac{\partial y^{(1)}}{\partial x^{(3)}} \right) \frac{\partial x^{(1)}}{\partial y^{(1)}} + \text{etc.}\end{aligned}$$

or

$$\overline{X}^r = \rho_s{}^r X^s \qquad (s \text{ umbral})$$

where

$$\rho_s{}^r \equiv \rho_1 \frac{\partial x^{(r)}}{\partial y^{(1)}} \frac{\partial y^{(1)}}{\partial x^{(s)}} + \rho_2 \frac{\partial x^{(r)}}{\partial y^{(2)}} \frac{\partial y^{(2)}}{\partial x^{(s)}} + \rho_3 \frac{\partial x^{(r)}}{\partial y^{(3)}} \frac{\partial y^{(3)}}{\partial x^{(s)}}$$

Now $\dfrac{\partial x^{(r)}}{\partial y^{(1)}}$ is a contravariant tensor and $\dfrac{\partial y^{(1)}}{\partial x^{(s)}}$ is a covariant tensor if we regard the y's as fixed and consider merely transformations on the x's so that $\rho_s{}^r$, being the sum of three mixed tensors, is actually, as the notation implies, a mixed tensor of rank two. It was in this geometrical way that Voigt introduced the idea which he called a tensor. The mixed tensor $\rho_s{}^r$ is completely specified by the three directions y and the scalar numbers ρ_1, ρ_2, ρ_3. If the mixed tensor is to be symmetric for every choice of ρ_1, ρ_2, ρ_3 we must have

$$\frac{\partial x^{(r)}}{\partial y^{(1)}} \frac{\partial y^{(1)}}{\partial x^{(s)}} \equiv \frac{\partial x^{(s)}}{\partial y^{(1)}} \cdot \frac{\partial y^{(1)}}{\partial x^{(r)}}, \qquad \text{etc.}$$

These equations lead to the conclusion that the "axes" y of

the tensor are mutually at right angles and so such a tensor was called symmetric.

In order to study the behavior of the vector $\overline{X}$ as X changes direction, keeping its magnitude unaltered, we may solve the equations for $\overline{X}$ and obtain

$$X^r \equiv \pi_s{}^r \overline{X}^s \qquad (s \text{ umbral})$$

where from the geometrical construction $\pi_s{}^r$ is a mixed tensor with the same axes as $\rho_s{}^r$ but

$$\pi_1 = \frac{1}{\rho_1}, \qquad \text{etc.,}$$

so that

$$\pi_s{}^r \equiv \frac{1}{\rho_1}\frac{\partial x^{(r)}}{\partial y^{(1)}}\frac{\partial y^{(1)}}{\partial x^{(s)}} + \frac{1}{\rho_2}\frac{\partial x^{(r)}}{\partial y^{(2)}}\frac{\partial y^{(2)}}{\partial x^{(s)}} + \frac{1}{\rho_3}\frac{\partial x^{(r)}}{\partial y^{(3)}}\frac{\partial y^{(3)}}{\partial x^{(s)}}.$$

Then squaring and adding the equations for X^r we find that $\overline{X}$ traces an ellipsoid, called the first tensor ellipsoid.

For a symmetric tensor the directions y are orthogonal so that $Y_1 = Y^1$, etc. A simple example of a symmetric tensor is furnished by the uniform stretching of a medium along three mutually perpendicular directions successively. It was from this example that Voigt originally took the name "Tensor." Reference may be made to any treatise on the Theory of Elasticity for an amplification of the remarks of this paragraph.

5. General Form of Green's Fundamental Lemma

Starting with any invariant function of position $V(x^{(1)} \cdots x^{(n)})$ we have seen how to form its covariant tensor gradient

$$X_r \equiv \frac{\partial V}{\partial x^{(r)}} \qquad (r = 1 \cdots n)$$

the square of whose magnitude is the *first differential parameter* of V

$$\Delta_1 V \equiv g^{rs}\frac{\partial V}{\partial x^{(r)}} \cdot \frac{\partial V}{\partial x^{(s)}}.$$

Now the normal direction tensor to $V(x^{(1)} \cdots x^{(n)}) =$ const. has components whose ratios are

$$n^{(1)} : n^{(2)} : \cdots : n^{(n)} = g^{1s}\frac{\partial V}{\partial x^{(s)}} : \cdots : g^{ns}\frac{\partial V}{\partial x^{(s)}} \qquad (s \text{ umbral})$$

the actual magnitudes of these being found on division through by the positive square root of $\Delta_1 V$. Hence the resolved part of the covariant tensor gradient along the normal is

$$\frac{\partial V}{\partial x^{(s)}} g^{st} \frac{\partial V}{\partial x^{(t)}} \div \sqrt{\Delta_1 V} \qquad (s,\ t \text{ umbral})$$

and this is $\equiv \sqrt{\Delta_1 V}$.* This we shall call the *normal derivative* of V and denote by the *symbol* $\dfrac{\partial V}{\partial n}\cdot$ The resolved part of the gradient along any *direction* l is

$$l^{(r)} \frac{\partial V}{\partial x^{(r)}} \qquad (r \text{ umbral})$$

This we denote by $\dfrac{\partial V}{\partial l}$ and *call the directional derivative* of V along the direction l. The angle θ between n and l is given by

$$\cos\theta = g_{rs} l^{(r)} n^{(s)} = \frac{1}{\sqrt{\Delta_1 V}} g_{rs} l^{(r)} g^{st} \frac{\partial V}{\partial x^{(t)}} \qquad (r,\ s,\ t \text{ umbral})$$

$$= \frac{1}{\sqrt{\Delta_1 V}} l^{(r)} \frac{\partial V}{\partial x^{(r)}} = \frac{1}{\sqrt{\Delta_1 V}} \frac{\partial V}{\partial l}$$

Hence

$$\frac{\partial V}{\partial l} \equiv \frac{\partial V}{\partial n} \cos\theta$$

showing that the maximum directional derivative is that along the normal. (In general, if we say that any covariant tensor X_r has a direction specified by the reciprocal contravariant tensor

$$X^s \equiv g^{sr} X_r \qquad (r \text{ umbral})$$

* If we define the "direction" of any covariant tensor of rank one as that of its reciprocal contravariant tensor we may say that the *gradient* of any invariant function of position has a direction normal to it.

the resolved part of X_r along any direction l is the product of the magnitude of the tensor into the cosine of the angle between l and the direction of the tensor.)

The contravariant tensor reciprocal to the gradient of V is

$$X^r \equiv g^{rs}\frac{\partial V}{\partial x^{(s)}}$$

Accordingly, on multiplying each of these expressions by $\sqrt{g}$, we derive the n distinct components of an alternating covariant tensor of rank $n-1$ (cf. Ch. 3, § 4) and so we can form the integral I_{n-1}

$$I_{n-1} \equiv \int \left\{\sqrt{g}\, g^{rs}\frac{\partial V}{\partial x^{(s)}}(J_r)\right\} du_1 \cdots du_{n-1}$$

over any spread of $n-1$ dimensions given by

$$x^{(s)} \equiv x^{(s)}(u_1, \cdots, u_{n-1}) \qquad (s = 1, \cdots, n)$$

the symbol (J_r) denoting as before

$$(-1)^{n-r}\frac{\partial(x^{(1)} \cdots x^{(r-1)}\, x^{(r+1)} \cdots x^{(n)})}{\partial(u_1 \cdots u_{n-1})}$$

The normal contravariant tensor to the spread of $n-1$ dimensions has, as has been shown, components proportional to

$$g^{rs}(J_s) \qquad (r = 1, \cdots, n;\ s \text{ umbral})$$

the actual magnitudes being found by dividing through by the positive square root of

$$g_{lm}g^{ls}(J_s)g^{mt}(J_t) \qquad (l, m, s, t \text{ umbral})$$
$$\equiv g^{st}(J_s)(J_t) \qquad (s, t \text{ umbral})$$

Hence $g^{rs}(J_s)\dfrac{\partial V}{\partial x^{(r)}} =$ product of $\sqrt{g^{st}(J_s)(J_t)}$by $\dfrac{\partial V}{\partial n}$ the directional derivative V normal to the spread V_{n-1} over which I_{n-1} is being extended. Hence we may write

$$I_{n-1} \equiv \int \frac{\partial V}{\partial n}\, dV_{n-1}$$

where by dV_{n-1} we mean the invariant $\sqrt{gg^{mt}(J_m)(J_t)}\ du_1 \cdots du_{n-1}$. (That this is invariant follows from rule (*d*) since $\sqrt{g}\,(J_r)$ is a covariant tensor of rank one (cf. Ch. 3, § 7).)

Applying Stokes' Lemma to I_{n-1} we have

$$\int \frac{\partial V}{\partial n} dV_{n-1} \equiv \int \Delta_2 V \cdot dV_n$$

where the integral I_{n-1} is extended over any V_{n-1} which is *closed* and the integral I_n on the right is extended over any region of space V_n bounded by V_{n-1}. Here

$$\Delta_2 V = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^s} \left(\sqrt{g}\, g^{rs} \frac{\partial V}{\partial x^r} \right)$$

and dV_n is the invariant $\sqrt{g}\ d(x^{(1)} \cdots x^{(n)})$.

If, instead of the contravariant tensor $X^r = g^{rs} \dfrac{\partial V}{\partial x^{(s)}}$, we start out with

$$X^r \equiv g^{rs}\, U \frac{\partial V}{\partial x^s}$$

where U is an invariant function of position we find

$$\int U \frac{\partial V}{\partial n} dV_{n-1} \equiv \int \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^s} \left(\sqrt{g}\, g^{rs}\, U \frac{\partial V}{\partial x^{(r)}} \right) dV_n$$

On interchanging the functions U, V and subtracting we have

$$\int \left(U \frac{\partial V}{\partial n} - V \frac{\partial U}{\partial n} \right) dV_{n-1} \equiv \int (U \Delta_2 V - V \Delta_2 U) dV_n$$

which is the usual form of Green's Lemma. The previous equation may be written

$$\int U \frac{\partial V}{\partial n} dV_{n-1} \equiv \int (U \Delta_2 V - \Delta(U, V)) \cdot dV_n$$

where $\Delta(U, V)$ is the invariant *mixed* differential parameter

$$\Delta(U, V) \equiv g^{rs} \frac{\partial V}{\partial x^{(r)}} \frac{\partial U}{\partial x^{(s)}} \qquad (r, s \text{ umbral})$$

In particular, if the invariant functions U, V are identical we have

$$\int U \frac{\partial U}{\partial n} \cdot dV_{n-1} \equiv \int (U\Delta_2 V - \Delta_1 U) \cdot dV_n$$

The last identity is the basis of various *uniqueness* theorems of Theoretical Physics. If we know the values of U or $\frac{\partial U}{\partial n}$ over a closed V_{n-1} as well as the values of $\Delta_2 U$ throughout the region bounded by V_{n-1} the function U is unique, save possibly to an unimportant additive constant. For, applying the last identity to the function $W \equiv U_1 - U_2$ where U_1 and U_2 satisfy the above conditions, we have

$$\int \Delta_1 W \cdot dV_n = 0$$

Now under the hypothesis that

$$(ds)^2 \equiv g_{rs} dx^{(r)} dx^{(s)} \equiv g^{rs} \xi_r \xi_s$$

is a definite form we see that $\Delta_1 W$ is one signed and vanishes only when all $\frac{\partial W}{\partial x^{(r)}}$ are zero. Hence since $\int \Delta_1 W \cdot dV_n = 0$ we must have $\frac{\partial W}{\partial x^{(r)}} \equiv 0$ throughout the region of integration $(r = 1 \cdots n)$. Therefore, W is a constant and if the values of U are assigned

$$W \equiv U_1 - U_2 = 0$$

on the boundary and so $W \equiv 0$ or $U_1 \equiv U_2$.

The whole argument depends on the definiteness of $(ds)^2$. Suppose we wish to apply the theorem to solutions of the *wave equation*

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 V}{\partial t^2} = 0$$

Here we have

$$(ds)^2 = (dx)^2 + (dy)^2 + (dz)^2 - c^2(dt)^2$$

and so

$$\Delta_1 V \equiv \left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2 - \frac{1}{c^2}\left(\frac{\partial V}{\partial t}\right)^2$$

and the theorem cannot be applied since $\Delta_1 V$ can vanish without implying the vanishing of all the derivatives.

6. Application to Maxwell's Equations

One of the most interesting applications of the algebra of tensors is the discussion of Maxwell's Electromagnetic Equations. These consist of two sets, which in the symbols of restricted vector analysis and the units employed by Heaviside are

$$(a) \quad -\frac{1}{c}\frac{\partial D}{\partial t} + \operatorname{curl} H = j; \quad \operatorname{div} D = \rho$$

$$(b) \quad \frac{1}{c}\frac{\partial B}{\partial t} + \operatorname{curl} E = 0; \quad \operatorname{div} B = 0$$

D is the electric *displacement,* H the magnetic force, and j the current vector; B is the magnetic *induction,* E the electric force and ρ is the volume density of electrification. We take $n = 4$ and as coordinates, in the above form,

$$x^{(1)} = x; \qquad x^{(2)} = y; \qquad x^{(3)} = z; \qquad x^{(4)} = t$$

If we *assume* that

$$X_{23} \equiv B_x; \quad X_{31} \equiv B_y; \quad X_{12} \equiv B_z; \quad X_{14} \equiv cE_x; \quad X_{24} \equiv cE_y; \quad X_{34} \equiv cE_z$$

are the six distinct components of an alternating covariant tensor of rank two, the four equations (b) express that

$$X_{123} \equiv \frac{\partial X_{12}}{\partial z} + \frac{\partial X_{23}}{\partial x} + \frac{\partial X_{31}}{\partial y} \equiv \frac{\partial B_z}{\partial z} + \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} \equiv \operatorname{div} B = 0,$$

$$X_{124} \equiv \frac{\partial X_{12}}{\partial t} + \frac{\partial X_{24}}{\partial x} + \frac{\partial X_{41}}{\partial y} \equiv \frac{\partial B_z}{\partial t} + c\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right) = 0,$$

$$X_{134} \equiv \frac{\partial X_{13}}{\partial t} + \frac{\partial X_{34}}{\partial x} + \frac{\partial X_{41}}{\partial z} \equiv -\frac{\partial B_y}{\partial t} + c\left(\frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z}\right) = 0,$$

$$X_{234} \equiv \frac{\partial X_{23}}{\partial t} + \frac{\partial X_{34}}{\partial y} + \frac{\partial X_{42}}{\partial z} \equiv \frac{\partial B_x}{\partial t} + c\left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) = 0$$

In other words the integral

$$
\begin{aligned}
I_2 &\equiv \int X_{rs} d(x^{(r)}, x^{(s)}) \\
&\equiv \int B_x d(y, z) + B_y d(z, x) + B_z d(x, y) + cE_x d(x, t) \\
&\qquad\qquad + cE_y d(y, t) + cE_z d(z, t)
\end{aligned}
$$

is the integral of an exact differential—its value when extended over any *closed* spread V_2 is identically zero. Hence its value when extended over any open spread V_2 can be expressed as a line integral $\int X_r dx^{(r)}$ round its boundary. On writing

$$X_1 \equiv -A_x; \qquad X_2 \equiv -A_y; \qquad X_3 \equiv -A_z; \qquad X_4 \equiv c\phi$$

we have

$$I_2 \equiv I_1 \equiv \int (c\phi dt - A_x dy - A_y dy - A_z dz)$$

and an application of Stokes' Lemma tells us that

$$X_{rs} \equiv \frac{\partial X_r}{\partial x^{(s)}} - \frac{\partial X_s}{\partial x^{(r)}}$$

or

$$
\begin{aligned}
B_x &\equiv X_{23} \equiv \frac{\partial X_2}{\partial x^{(3)}} - \frac{\partial X_3}{\partial x^{(2)}} \equiv \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}; \\
B_y &\equiv \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}; \qquad B_z \equiv \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}; \\
cE_x &\equiv X_{14} \equiv \frac{\partial X_1}{\partial x^{(4)}} - \frac{\partial X_4}{\partial x^{(1)}} \equiv -\frac{\partial A_x}{\partial t} - c\frac{\partial \phi}{\partial x}; \\
cE_y &\equiv -\frac{\partial A_y}{\partial t} - c\frac{\partial \phi}{\partial y}; \qquad cE_z \equiv -\frac{\partial A_z}{\partial t} - c\frac{\partial \phi}{\partial z}
\end{aligned}
$$

The covariant tensor of the first rank $(A_x, A_y, A_z, -c\phi)$ is the "electromagnetic covariant tensor potential" of which the first three components form Maxwell's *vector potential,* ϕ being his scalar potential.

Similarly, if we *assume* that $(-D_x, -D_y, -D_z, cH_x, cH_y, cH_z)$ are the six distinct functions of an alternating covariant tensor $\overline{X}_{rs}$ of rank two the equations (a) say that

$$\overline{X}_{234} \equiv cj_x; \qquad \overline{X}_{314} \equiv cj_y; \qquad \overline{X}_{124} \equiv cj_z; \qquad \overline{X}_{123} = -\rho$$

and we have $I_2 \equiv I_3$ where

$$I_2 \equiv \int cH_x d(x, t) + cH_y d(y, t) + cH_z d(z, t) - D_x d(y, z) - D_y d(z, x) - D_z d(x, y),$$

$$I_3 \equiv \int cj_x d(y, z, t) + cj_y d(z, x, t) + cj_z d(x, y, t) - \rho d(x, y, z)$$

I_2 being taken over any closed spread V_2 of two dimensions and I_3 being taken over the open V_3 bounded by V_2. Accordingly $(j_x, j_y, j_z, -\rho/c)$ are the four distinct functions of an alternating tensor of rank three and so, on writing $c(\overline{X}_1) \equiv X_{234}$, etc., $\dfrac{(\overline{X}_r)}{\sqrt{g}}$ form a contravariant tensor of rank one (Ch. 3, § 7). From its definition we know that its *divergence is zero*. This tensor we may call the current contravariant tensor and write

$$C^1 = \frac{j_1}{\sqrt{g}}; \quad \cdots \quad C^4 = \frac{-\rho}{c\sqrt{g}}$$

Let us now apply these methods to the problem of writing Maxwell's equations in a form suitable for work with curvilinear coordinates $y^{(1)}, y^{(2)}, y^{(3)}$ in space of three dimensions—the time t not entering into the transformation. The equations connecting the x and y coordinates are of the type

$$x^{(1)} \equiv x^{(1)}(y^{(1)}, y^{(2)}, y^{(3)}); \qquad x^{(2)} \equiv x^{(2)}(y^{(1)}, y^{(2)}, y^{(3)});$$
$$x^{(3)} \equiv x^{(3)}(y^{(1)}, y^{(2)}, y^{(3)}); \qquad x^{(4)} \equiv y^{(4)} = t$$

and denoting tensor components in the new coordinate system by primes we have

$$(D_1)' \equiv D_{23}' \equiv D_{23}\frac{\partial(x^{(2)}, x^{(3)})}{\partial(y^{(2)}, y^{(3)})} + D_{31}\frac{\partial(x^{(3)}, x^{(1)})}{\partial(y^{(2)}, y^{(3)})} + D_{12}\frac{\partial(x^{(1)}, x^{(2)})}{\partial(y^{(2)}, y^{(3)})}$$

the terms in H_1, H_2, H_3 vanishing since

$$\frac{\partial x^{(4)}}{\partial y^{(2)}} \equiv 0 \equiv \frac{\partial x^{(4)}}{\partial y^{(3)}}$$

$$(H_1)' \equiv (H_1)\frac{\partial(x^{(1)}, x^{(4)})}{\partial(y^{(1)}, y^{(4)})} + (H_2)\frac{\partial(x^{(2)}, x^{(4)})}{\partial(y^{(1)}, y^{(4)})} + (H_3)\frac{\partial(x^{(3)}, x^{(4)})}{\partial(y^{(1)}, y^{(4)})}$$
$$\equiv (H_1)\frac{\partial x^{(1)}}{\partial y^{(1)}} + (H_2)\frac{\partial x^{(2)}}{\partial y^{(1)}} + (H_3)\frac{\partial x^{(3)}}{\partial y^{(1)}}$$

the terms in (D_1) (D_2) (D_3) vanishing since

$$\frac{\partial x^{(1)}}{\partial y^{(4)}} \equiv \frac{\partial x^{(2)}}{\partial y^{(4)}} \equiv \frac{\partial x^{(3)}}{\partial y^{(4)}} \equiv 0$$

Hence in the three-dimensional space with coordinate systems $(x^{(1)}, x^{(2)}, x^{(3)})$ and $(y^{(1)}, y^{(2)}, y^{(3)})$ the variable t being regarded merely as a parameter which does not enter into

$$(ds)^2 \equiv g_{rs}dx^{(r)}dx^{(s)} \equiv f_{rs}dy^{(r)}dy^{(s)} \quad (r, s = 1, 2, 3)$$

the three quantities (D_1) (D_2) (D_3) are the three distinct members of an alternating covariant tensor of rank two. Hence $\frac{(D_r)}{\sqrt{g}} \equiv X^r$ is a contravariant tensor of rank one; similarly $\frac{(B_r)}{\sqrt{g}} \equiv \overline{X}^r$ is a contravariant tensor of rank one whilst $E_r \equiv X_r$ and $H_r \equiv \overline{X}_r$ are *covariant* tensors of rank one. We derive by our rule (*d*) of composition the invariants

$$\frac{(ED)}{\sqrt{g}}; \qquad \frac{(EB)}{\sqrt{g}}; \qquad \frac{(HD)}{\sqrt{g}}; \qquad \frac{(HB)}{\sqrt{g}}$$

where as in the usual vector notation

$$(ED) \equiv E_1D_1 + E_2D_2 + E_3D_3$$

and similarly for the others.

Dividing Maxwell's equations, as usually written, across by $\sqrt{g}$ we obtain

$$-\frac{1}{c}\frac{\partial}{\partial t}X^r + \frac{1}{\sqrt{g}}\operatorname{curl}_r(H) = C^r \qquad (r = 1, 2, 3)$$

(where $C^r \equiv \frac{\dot{\jmath}_r}{\sqrt{g}}$ is the contravariant current vector).

$$\operatorname{div} X^{(r)} = \rho$$

where ρ is the invariant charge density and similarly from the second set

$$+\frac{1}{c}\frac{\partial}{\partial t}\overline{X}^r + \frac{1}{\sqrt{g}}\operatorname{curl}_r (E) = 0$$

$$\operatorname{div} \overline{X}^r = 0$$

Denoting, then, as usual resolved parts along the coordinate lines by subscripts (l_1, l_2, l_3) and along the normals to the coordinate surfaces by the subscripts (n_1, n_2, n_3) we have the three equations

$$-\frac{1}{c}\sqrt{f^{11}}\frac{\partial}{\partial t}D_{n_1}+\frac{1}{\sqrt{f}}\left\{\frac{\partial}{\partial y^{(2)}}(\sqrt{f_{33}}\,H_{l_3})-\frac{\partial}{\partial y^{(3)}}(\sqrt{f_{22}}\,H_{l_2})\right\}=\sqrt{f^{11}}\,C_{n_1},$$

$$-\frac{1}{c}\sqrt{f^{22}}\frac{\partial}{\partial t}D_{n_2}+\frac{1}{\sqrt{f}}\left\{\frac{\partial}{\partial y^{(3)}}(\sqrt{f_{11}}\,H_{l_1})-\frac{\partial}{\partial y^{(1)}}(\sqrt{f_{33}}\,H_{l_3})\right\}=\sqrt{f^{22}}\,C_{n_2},$$

$$-\frac{1}{c}\sqrt{f^{33}}\frac{\partial}{\partial t}D_{n_3}+\frac{1}{\sqrt{f}}\left\{\frac{\partial}{\partial y^{(1)}}(\sqrt{f_{22}}\,H_{l_2})-\frac{\partial}{\partial y^{(2)}}(\sqrt{f_{11}}\,H_{l_1})\right\}=\sqrt{f^{33}}\,C_{n_3}$$

The equation $\operatorname{div} X^r = \rho$ becomes

$$\frac{1}{\sqrt{f}}\left\{\frac{\partial}{\partial y^{(1)}}(\sqrt{ff^{11}}\,D_{n_1})+\frac{\partial}{\partial y^{(2)}}(\sqrt{ff^{22}}\,D_{n_2})+\frac{\partial}{\partial y^{(3)}}(\sqrt{ff^{33}}\,D_{n_3})\right\}=\rho$$

(by D_{n_1} is meant the resolved part of the contravariant tensor $D/\sqrt{g}$ along the direction n_1).

The equations (b) are similar and are simplified by the fact that there C_{n_1}, C_{n_2}, C_{n_3}, ρ are all zero.*

* When the coordinates y are orthogonal

$$(ds)^2 = \frac{1}{h_1^2}(dy^{(1)})^2 + \frac{1}{h_2^2}(dy^{(2)})^2 + \frac{1}{h_3^2}(dy^{(3)})^2$$

$f = \dfrac{1}{h_1^2h_2^2h_3^2}$; $f^{11} = h_1^2$, etc., and Maxwell's equations become since $n_1 = l_1$, etc.

$$-\frac{1}{c}\frac{\partial}{\partial t}D_{l_1} + h_2h_3\left\{\frac{\partial}{\partial y^{(2)}}\left(\frac{H_{l_3}}{h_3}\right) - \frac{\partial}{\partial y^{(3)}}\left(\frac{H_{l_2}}{h_2}\right)\right\} = C_{l_1}$$

and two similar equations together with

$$h_1h_2h_3\left\{\frac{\partial}{\partial y^{(1)}}\left(\frac{D_{l_1}}{h_2h_3}\right)+\frac{\partial}{\partial y^{(2)}}\left(\frac{D_{l_2}}{h_3h_1}\right)+\frac{\partial}{\partial y^{(3)}}\left(\frac{D_{l_3}}{h_1h_2}\right)\right\} = \rho.$$

Cf. *Riemann-Weber*, *Die Partiellen Differentialgleichungen der Mathematischen Physik*, Bd. 2, p. 312 (Vieweg & Sohn) (1919).

EXAMPLE

In space polar coordinates Maxwell's equations are

$$-\frac{1}{c}\frac{\partial}{\partial t}D_r + \frac{1}{r^2 \sin\theta}\left\{\frac{\partial}{\partial\theta}(r\sin\theta H_\phi) - \frac{\partial}{\partial\phi}(rH_\theta)\right\} = C_r$$

$$-\frac{1}{c}\frac{\partial}{\partial t}D_\theta + \frac{1}{r\sin\theta}\left\{\frac{\partial}{\partial\phi}(H_r) - \frac{\partial}{\partial r}(r\sin\theta H_\phi)\right\} = C_\theta$$

$$-\frac{1}{c}\frac{\partial}{\partial t}D_\phi + \frac{1}{r}\left\{\frac{\partial}{\partial r}(rH_\theta) - \frac{\partial}{\partial\theta}(H_r)\right\} = C_\phi$$

$$\frac{1}{r^2\sin\theta}\left\{\frac{\partial}{\partial r}(r\sin\theta D_r) + \frac{\partial}{\partial\theta}(r\sin\theta D_\theta) + \frac{\partial}{\partial\phi}(rD_\phi)\right\} = \rho$$

It is particularly to be noticed that Maxwell's Equations are essentially of a non-metrical character. No real simplification is introduced by the hypothesis that the fundamental space is of the ordinary Euclidean character. Another point to which attention should be directed is the difference in character of the tensors B and H or of D and E. A relation of the familiar type

$$B = \mu H$$

μ, the coefficient of permeability, being supposed invariant is not the proper mode of statement of a physical law if we understand by B_1, B_2, B_3 the three components of the tensor B. The true statement of the law is

$$(B)_l = \mu\cdot(H)_l$$

where by $(B)_l$ we mean the resolved part of the contravariant tensor $(B)/\sqrt{g}$ along the direction l and by $(H)_l$ we mean the resolved part of the covariant tensor H along the same direction. Thus any *constitutive equation* of this type is an allowable statement of a physical law *not* because it is a *tensor* equation (since it is not such), but because it is an equality between invariant magnitudes or a *scalar* equation. The true tensor equation is found by equating the covariant tensor μH to the covariant tensor reciprocal to the contravariant tensor $(B)/\sqrt{g}$.

CHAPTER V

1. Connection of Tensor Algebra with Integral Invariants and Application to the Statement of Faraday's Law of Moving Circuits*

Suppose for example we have a curve V_1 whose equations

$$x^{(s)} \equiv x^{(s)}(u, \tau) \qquad (s = 1, \cdots, n)$$

involve a *parameter* τ. This curve may be said, adopting the language of dynamics, to move and trace out a V_2 whose equations are those given above, the parameters being u and τ. Any one of the curves $\tau =$ constant will then be a position of the moving curve. We shall suppose that the values of u serve to identify the various points on the moving curve; thus if u denotes the distance along the initial position of the moving curve from a certain fixed point, or origin, the curves $\overline{V}_1$ obtained by taking $u =$ constant (u_0) in the equations

$$x^{(s)} \equiv x^{(s)}(u, \tau) \qquad (s = 1, \cdots, n)$$

are the *path curves* of the definite point on the curve V_1 which initially was at the distance u_0 from the origin on V_1. It will fix our ideas to consider V_1 as made up of particles of a fluid; then the curves $\overline{V}_1$ are the paths of the various *material particles* of V_1. It is well to insist, at the outset, on the point that the parameters u and τ are independent. Thus if the moving curve V_1 were rigid, u could be taken as the arc distance along V_1 at

* An elementary presentation of the theory of Integral Invariants is given by Goursat, E., in two papers:

(*a*) Sur les invariants intégraux. Journal de Mathématiques, 6e série, t. IV (1908), p. 331.

(*b*) Sur quelques points de la théorie des invariants intégraux. Journal de mathématiques, 7e série, t. 1 (1915), p. 241.

any time τ; if, however, as in the case of the curve made up of material *fluid* particles, V_1 is not rigid, u may only be taken as the *initial* identifying arc distance; otherwise u would vary with τ. Let us now consider an integral $I_1 \equiv \int X_r dx^{(r)}$ extended over V_1 and ask the conditions that I_1 should be the same for all the curves V_1, i.e., that I_1 should not vary with τ. If this is so, I_1 is said to be an *integral invariant.*

Now I_1 is in general a function of τ defined by

$$I_1(\tau) \equiv \int_{u_0}^{u'} \left(X_r \frac{\partial x^{(r)}}{\partial u} \right) du \qquad (r \text{ umbral})$$

the limits u_0 and u' being, however, since u and τ are independent, quite independent of τ. Hence

$$\frac{dI_1}{d\tau} \equiv \int_{u_0}^{u'} \frac{d}{d\tau} \left(X_r \frac{\partial x^{(r)}}{\partial u} \right) du$$

The coefficients X_r are functions of position and therefore involve τ indirectly; it is somewhat more general to contemplate the possibility that they may involve τ, not only in this indirect manner but also directly. Then for any one of the coefficients X_r we have

$$\frac{d}{d\tau} X_r \equiv \left(\frac{\partial X_r}{\partial \tau} + \frac{\partial X_r}{\partial x^{(s)}} \frac{\partial x^{(s)}}{\partial \tau} \right)$$

It is now convenient to denote the contravariant tensor of rank one $-\dfrac{\partial x^{(r)}}{\partial \tau}-$ by the symbol X^r and to use the result

$$\frac{d}{d\tau} \frac{\partial x^{(r)}}{\partial u} \equiv \frac{\partial}{\partial u} \frac{\partial x^{(r)}}{\partial \tau} \equiv \frac{\partial X^{(r)}}{\partial u} \equiv \frac{\partial X^{(r)}}{\partial x^{(s)}} \frac{\partial x^{(s)}}{\partial u} \qquad (s \text{ umbral})$$

and we have

$$\frac{dI_1}{d\tau} \equiv \int \left(\frac{dX_r}{d\tau} \frac{\partial x^{(r)}}{\partial u} + X_r \frac{d}{d\tau} \frac{\partial x^{(r)}}{\partial u} \right) du \qquad (r \text{ umbral})$$

$$\equiv \int \left\{ \left(\frac{\partial X_r}{\partial \tau} + X^{(s)} \frac{\partial X_r}{\partial x^{(s)}} \right) \frac{\partial x^{(r)}}{\partial u} + X_r \frac{\partial X^{(r)}}{\partial x^{(s)}} \frac{\partial x^{(s)}}{\partial u} \right\} du$$

$$(r, s \text{ umbral})$$

$$= \int \left\{ \frac{\partial X_r}{\partial \tau} + X^s \frac{\partial X_r}{\partial x^{(s)}} + X_s \frac{\partial X^{(s)}}{\partial x^{(r)}} \right\} \frac{\partial x^{(r)}}{\partial u} du$$

(on modifying suitably the umbral symbols)

Hence if $dI_1/d\tau$ is to vanish identically for all curves V_1 we must have

$$\frac{\partial X_r}{\partial t} + X^s \frac{\partial X_r}{\partial x^{(s)}} + X_s \frac{\partial X^{(s)}}{\partial x^{(r)}} \equiv 0 \quad (r = 1, \cdots, n, s \text{ umbral}).$$

Sometimes it is only necessary that I_1 should be unchanged for all *closed* curves V_1; in this case I_1 is said to be a *relative* integral invariant. To find the conditions for this we use Stokes' Lemma to replace the I_1 over a *closed* curve by an I_2 over an open V_2 and then find the conditions that I_2 should be an (absolute) integral invariant.

The analysis necessary to find the conditions that an

$$I_p \equiv \int X_{s_1 \cdots s_p} d(x^{(s_1)} \cdots x^{(s_p)})$$

extended over a V_p (moving) whose equations are

$$x^{(s)} \equiv x^{(s)}(u_1 \cdots u_p, \tau) \qquad (s = 1 \cdots n)$$

should be an absolute invariant is identical with that given for the simplest case $p = 1$. Let us write as before

$$\frac{\partial x^{(s)}}{\partial \tau} \equiv X^{(s)}$$

and denote by the symbol $\dot{F}$ the derivative

$$\frac{dF}{d\tau} \equiv \frac{\partial F}{\partial \tau} + \frac{\partial F}{\partial x^{(s)}} X^{(s)} \qquad (s \text{ umbral})$$

where F is any function of position which may also involve τ explicitly. Then

$$\frac{dI_p}{d\tau} \equiv \int \frac{d}{d\tau} \{X_{s_1 \cdots s_p} d(x^{(s_1)} \cdots x^{(s_p)})\} \quad (s_1 < s_2 < \cdots \text{ umbral})$$

since the limits of integration with respect to the variables u are independent of τ. This we write

$$\equiv \int \frac{d}{d\tau}\left\{X_{s_1 \cdots s_p}\frac{\partial x^{(s_1)}}{\partial u_1}\cdots\frac{\partial x^{(s_p)}}{\partial u_p}\right\} du_1 \cdots du_p$$

and availing ourselves of the relation

$$\frac{d}{d\tau}\frac{\partial x^{(s_r)}}{\partial u_r} = \frac{\partial}{\partial u_r}X^{(s_r)} \equiv \frac{\partial X^{(s_r)}}{\partial x^{(m)}}\frac{\partial x^{(m)}}{\partial u_r}$$

we arrive at the conditions expressed in the form that

$$\dot{X}_{s_1 \cdots s_p} + X_{ms_2 \cdots s_p}\frac{\partial X^{(m)}}{\partial x^{(s_1)}} + X_{s_1 m s_3 \cdots s_p}\frac{\partial X^{(m)}}{\partial x^{(s_2)}} + \cdots$$
$$+ X_{s_1 \cdots s_{p-1} m}\frac{\partial X^{(m)}}{\partial x^{(s_p)}} \equiv 0 \quad (m \text{ umbral})$$

An especially simple case is that in which $p = n$. Here there is a single condition

$$\dot{X}_{1 \cdots n} + X_{1 \cdots n}\left(\frac{\partial X^{(r)}}{\partial x^{(r)}}\right) \qquad (r \text{ umbral})$$

Since $X_{1 \cdots n}$ is the single distinct member of an alternating covariant tensor of rank n

$$X_{1 \cdots n} = \sqrt{g}\, U$$

where U is an invariant function of position and writing out

$$\dot{X}_{1 \cdots n} \equiv \frac{\partial X_{1 \cdots n}}{\partial \tau} + X^s \frac{\partial X_{1 \cdots n}}{\partial x^{(s)}}$$

our condition that $\int U \cdot dV_n$ should be an integral invariant may be written in the form

$$\frac{\partial}{\partial \tau}(\sqrt{g}\, U) + \frac{\partial}{\partial x^{(s)}}(\sqrt{g}\ U \cdot X^{(s)}) = 0$$

or on dividing out by $\sqrt{g}$, which does not involve τ explicitly, $\frac{\partial U}{\partial \tau} + \operatorname{div}(UX^{(r)})$ where as usual the divergence of the contra-

variant tensor of rank one UX^r is the invariant $\frac{1}{\sqrt{g}}\frac{\partial}{\partial x^{(r)}}(\sqrt{g}\, UX^r)$. In this form the invariance of the condition for an integral invariant is apparent. If we are considering a moving charged material body where ρ is the density of charge, the *total charge* $\int \rho dV_n$ *remaining constant* gives us that

$$\frac{\partial \rho}{\partial t} + \text{div}\,(\rho X^{(r)}) = 0$$

where $X^{(r)}$ is the contravariant velocity tensor of rank one.

Faraday's Law for a Moving Circuit.

We have seen that

$$\frac{d}{d\tau}\int X_r dx^{(r)} \equiv \int \left\{\frac{\partial X_r}{\partial \tau} + X^s \frac{\partial X_r}{\partial x^{(s)}} + X_s \frac{\partial X^{(s)}}{\partial x^{(r)}}\right\} dx^{(r)}$$

the integral in each case being taken over the position of the moving curve at time τ. The expressions

$$\frac{\partial X_r}{\partial \tau} + X^s \frac{\partial X_r}{\partial x^{(s)}} + X_s \frac{\partial X^{(s)}}{\partial x^{(r)}} \quad (r = 1, \cdots, n;\ s \text{ umbral})$$

must accordingly form a covariant tensor of rank one. In fact we may write this as

$$\equiv \frac{\partial X_r}{\partial \tau} + X^s \left\{\frac{\partial X_r}{\partial x^{(s)}} - \frac{\partial X_s}{\partial x^{(r)}}\right\} + \frac{\partial}{\partial x^{(r)}}\{X_s X^{(s)}\}$$

when the covariant character is apparent by rule (*d*), Ch. 2, § 4, since

$$\frac{\partial X_r}{\partial x^{(s)}} - \frac{\partial X_{(s)}}{\partial x^{(r)}} \equiv X_{rs}$$

is covariant of rank two and $X_s X^{(s)}$ is invariant.

Let us now write down the expression for $dI_2/d\tau$ where I_2 is any surface integral and transform the coefficients as above so as to make evident their tensor character.

Writing

$$I_2 \equiv \int X_{rs} d(x^{(r)}, x^{(s)})$$

we get

$$\frac{dI_2}{d\tau} \equiv \int \overline{X}_{rs} d(x^{(r)}, x^{(s)})$$

where

$$\overline{X}_{rs} \equiv \frac{\partial X_{rs}}{\partial \tau} + X^t \frac{\partial X_{rs}}{\partial x^{(t)}} + X_{ms} \frac{\partial X^m}{\partial x^{(r)}} + X_{rm} \frac{\partial X^m}{\partial x^{(s)}}$$

the integrals being in each case extended over the position of the moving spread or surface V_2 at time τ. We may write

$$\overline{X}_{rs} \equiv \frac{\partial X_{rs}}{\partial \tau} + X^{(t)} \left\{ \frac{\partial X_{rs}}{\partial x^{(t)}} + \frac{\partial X_{st}}{\partial x^{(r)}} + \frac{\partial X_{tr}}{\partial x^{(s)}} \right\} + \left\{ \frac{\partial}{\partial x^{(s)}} (X_{rm} X^m) - \frac{\partial}{\partial x^{(r)}} (X_{sm} X^m) \right\}$$

where we have availed ourselves of the alternating character of X_{rs}. The covariant character of $\overline{X}_{rs}$ then follows from rule (d).

We shall apply this result to the surface integral

$$I_2 \equiv \int \{(D_1)d(x^{(2)}, x^{(3)}) + (D_2)d(x^{(3)}, x^{(1)}) + (D_3)d(x^{(1)}, x^{(2)})\} \qquad n = 3$$

so that (D_1), (D_2), (D_3) are the three distinct members of an alternating covariant tensor of rank two. Hence $D^{(r)} \equiv (D_r)/\sqrt{g}$ is a contravariant tensor of rank one. The covariant tensor of rank one whose curl appears in the expression for $\overline{X}_{rs}$ is

$$X_{rm} X^m \qquad (m \text{ umbral})$$

so that its first component is

$$\begin{aligned} X_{12} X^{(2)} + X_{13} X^{(3)} &\equiv X^{(2)}(D_3) - X^{(3)}(D_2) \\ &\equiv \{X^{(2)} D^{(3)} - X^{(3)} D^{(2)}\} \sqrt{g} \end{aligned}$$

It accordingly appears as that derived from the outer product of the velocity contravariant tensor and the displacement contravariant tensor.

The expression

$$\frac{\partial X_{rs}}{\partial x^{(t)}} + \frac{\partial X_{st}}{\partial x^{(r)}} + \frac{\partial X_{tr}}{\partial x^{(s)}} = \sqrt{g}\,\text{div}\,D^{(r)}$$

If now we *assume* as Maxwell's equations for the moving material medium

$$\frac{\partial}{\partial t}(D) = c\,\text{curl}\,H - (j);\ \ \text{div}\,D^r = \rho$$

where (j) is the alternating covariant current tensor of rank two, so that $(j)/\sqrt{g}$ is the contravariant current tensor of rank one C^r, we have for $\overline{X}_{rs}$ the equations

$$\overline{X}_{23} = \sqrt{g}\frac{\partial D^1}{\partial t} + X^{(1)}\sqrt{g}\,\rho + \left\{\frac{\partial}{\partial x^{(3)}}\sqrt{g}(X^{(3)}D^1 - X^1D^3) - \frac{\partial}{\partial x^{(2)}}\sqrt{g}(X^1D^2 - X^2D^1)\right\},$$

etc.

Using Stokes' Lemma to transform the surface integral of the part in face brackets into a line integral as well as that involving curl H in $\partial D^{(r)}/\partial t$ we find

$$\begin{aligned}\frac{d}{dt}\int D_n dS = \int \{cH_1 &- \sqrt{g}(X^{(2)}D^{(3)} - X^3D^2)\}dx_1 \\ &+ \{cH_2 - \sqrt{g}(X^3D^1 - X^1D^3)\}dx_2 \\ &+ \{cH_3 - \sqrt{g}(X^1D^2 - X^2D^1)\}dx_3 \\ &+ \int[\sqrt{g}(\rho X^r - C^r)d(x^s, x^t)]\end{aligned}$$

The integrand in the surface integral on the right is found by writing r, s, t in cyclic order and summing the terms corresponding to $r = 1, 2, 3$ respectively. (The line integral is to be taken over the boundary of the moving surface.) The contravariant tensor $\rho X^{(r)}$ is called the convection current. In exactly the same way we obtain, on making a similar assumption as to what Maxwell's equations should be for moving media,

$$\begin{aligned}-\frac{d}{dt}\int B_n dS \equiv \int\{cE_1 &+ \sqrt{g}(X^2B^3 - X^3B^2)\}dx_1 \\ &+ \{cE_2 + \sqrt{g}(X^3B^1 - X^1B^3)\}dx_2 \\ &+ \{cE_3 + \sqrt{g}(X^1B^2 - X^2B^1)\}dx_3\end{aligned}$$

there being now, however, no surface integral on the right-hand side. Accordingly the covariant tensor

$$E_r + \frac{1}{c}\sqrt{g}(X^{(s)}B^{(t)} - X^{(t)}B^{(s)}) \quad (r = 1, 2, 3;\ r, s, t \text{ cyclic})$$

is taken as the effective electric intensity along the moving curve; its line integral being called the effective electromotive force round the curve. ($X^{(r)}$ is the contravariant velocity tensor.) On multiplication by charge this tensor gives the *mechanical force* tensor.

Example. In space polar coordinates the mechanical force tensor per element of length on a moving curve with linear density σ is

$$\left\{E_r + \frac{1}{c}r^2 \sin\theta(v_\theta B_\phi - v_\phi B_\theta)\frac{1}{r^2\sin\theta}\right\}\sigma ds$$
$$\left\{rE_\theta + \frac{1}{c}r^2 \sin\theta(v_\phi B_r - v_r B_\phi)\frac{1}{r\sin\theta}\right\}\sigma ds$$
$$\left\{r\sin\theta E_\phi + \frac{1}{c}r^2 \sin\theta(v_r B_\theta - v_\theta B_r)\frac{1}{r}\right\}\sigma ds$$

where E_r, B_r, v_r are the *resolved* parts of E, B, X along the direction r and so on. To get the resolved parts of the mechanical force along the three coordinate directions multiply these by 1, $\frac{1}{r}$, $\frac{1}{r\sin\theta}$ respectively and we obtain the well-known formula

$$F = E + \frac{1}{c}[vB]$$

In the general case when the coordinates y_1, y_2, y_3 are not orthogonal the three resolved parts of the mechanical intensity (covariant) tensor along the coordinate lines y_1, y_2, y_3 respectively are

$$F_{l_1} \equiv E_{l_1} + \frac{1}{c}\sqrt{\frac{ff^{22}f^{33}}{f_{11}}}[v_{n_2}B_{n_3} - v_{n_3}B_{n_1}], \quad \text{etc.},$$

where v_{n_1} v_{n_2} v_{n_3} denote the resolved parts of the velocity along the normals to the coordinate surfaces $y_1 = \text{const.}$, $y_2 = \text{const.}$, $y_3 = \text{const.}$, respectively.

CHAPTER VI

1. The Tensor or Absolute Differential Calculus

Since the Calculus of Variations deals with properties of curves and surfaces without making any particular reference to the special coordinates used in describing the curves there must be a close relationship between that subject and that which we are discussing. It is this absolute or tensor character of the calculus of variations that has urged writers on Theoretical Physics to express the laws of physics, as far as possible, in the language of the Calculus of Variations. However, this subject has been placed on a clear and firm basis only within the past few decades and so it may be well to discuss one of its simpler problems—the more so as the solution of this problem is involved in the statement of Einstein's fundamental law of Inertia in the Theory of Relativity.

Let us consider a curve V_1, in space S_n of n dimensions, given by the equations

$$x^{(s)} \equiv x^{(s)}(u) \qquad (s = 1 \cdots n)$$

and in connection with this curve a function, *not merely of position,* but of the coordinates x and their *derivatives*

$$x^{(s)\prime} \equiv \frac{dx^{(s)}}{du}$$

The integral I_1 over the curve V_1 where

$$I_1 \equiv \int_{u_0}^{u'} F(x^{(1)} \cdots x^{(n)} x^{(1)\prime} \cdots x^{(n)\prime})du$$

has a value depending on the curve V_1 as well as on the particular function. The problem we wish to discuss is: What, if any, are the curves V_1 making, for a given function F, I_1 a minimum, all the curves V_1 being supposed to have the same end points.

To answer this question we consider a new curve $V_1(\alpha)$ given by the equations

$$x^{(s)} \equiv x^{(s)}(u, \alpha) \qquad (s = 1 \cdots n)$$

where α is quite independent of u. We suppose this parameter α to be such that when $\alpha = 0$, $V_1(\alpha)$ makes I_1 a minimum. $V_1(\alpha)$ is now completely determined by the equations just written when α is given and so I_1 is a function of (α) which may, we suppose, be expanded by Taylor's Theorem in the form

$$I(\alpha) = I(0) + \alpha\left(\frac{\partial I}{\partial \alpha}\right)_{\alpha=0} + \frac{\alpha^2}{1\cdot 2}\left(\frac{\partial^2 I}{\partial \alpha^2}\right)_{\alpha=0} + \cdots$$

This is written

$$I(\alpha) = I(0) + \delta I + \delta^2 I + \cdots$$

and δI is called the *first variation of the integral.* If $I(0)$ is to be a minimum it is necessary (although not always sufficient) that $\delta I = 0$ for otherwise $\Delta I \equiv I(\alpha) - I(0)$ would change sign with α when α is sufficiently small. Now the limits of the integral for I_1 are fixed and so to find $\partial I/\partial \alpha$ we have merely to differentiate the integrand F with respect to α. F involves α, not directly, but indirectly through the coordinates x and their derivatives x'.

Thus

$$\frac{\partial F}{\partial \alpha} = \frac{\partial F}{\partial x^{(s)}}\frac{\partial x^{(s)}}{\partial \alpha} + \frac{\partial F}{\partial x^{(s)\prime}}\frac{\partial x^{(s)\prime}}{\partial \alpha} \qquad (s \text{ umbral})$$

and therefore

$$\frac{\partial I}{\partial \alpha} = \int_{u_0}^{u'}\left(\frac{\partial F}{\partial x^{(s)}}\frac{\partial x^{(s)}}{\partial \alpha} + \frac{\partial F}{\partial x^{(s)\prime}}\frac{\partial x^{(s)\prime}}{\partial \alpha}\right)du$$

Now

$$\frac{\partial x^{(s)\prime}}{\partial \alpha} \equiv \frac{\partial^2 x^{(s)}}{\partial \alpha \partial u} = \frac{\partial}{\partial u}\cdot\frac{\partial x^{(s)}}{\partial \alpha}$$

so that, on integration by parts,

$$\int_{u_0}^{u'}\frac{\partial F}{\partial x^{(s)\prime}}\frac{\partial x^{(s)\prime}}{\partial \alpha}du = \frac{\partial F}{\partial x^{(s)\prime}}\frac{\partial x^{(s)}}{\partial \alpha}\Bigg|_{u_0}^{u'} - \int_{u_0}^{u'}\frac{\partial x^{(s)}}{\partial \alpha}\cdot\frac{\partial}{\partial u}\frac{\partial F}{\partial x^{(s)\prime}}du$$

Since the end points of the curve are *fixed*, $\partial x^{(s)}/\partial\alpha = 0$ at the limits of integration and so the integrated part vanishes and, collecting terms, we have

$$\frac{\partial I}{\partial \alpha} = \int_{u_0}^{u'} \left\{ \frac{\partial F}{\partial x^{(s)}} - \frac{\partial}{\partial u}\frac{\partial F}{\partial x^{(s)'}} \right\} \frac{\partial x^{(s)}}{\partial u}\, du \qquad (s \text{ umbral})$$

If $\left(\dfrac{\partial I}{\partial \alpha}\right)_{\alpha=0}$ is to be zero for *all possible* varied curves $V(\alpha)$ it is evidently sufficient and can be shown to be necessary that all the coefficients $\left(\dfrac{\partial F}{\partial x^{(s)}} - \dfrac{\partial}{\partial u}\dfrac{\partial F}{\partial x^{(s)'}}\right)_{\alpha=0}$ in this integral should vanish $(s = 1, \cdots, n)$.

These n expressions are the components of a covariant tensor of rank one where now, however, the term is used in a wider sense than hitherto. F is now not merely a function of the coordinates x but of their derivatives x'. From

$$x^{(r)'} \equiv \frac{\partial x^{(r)}}{\partial y^{(t)}} y^{(t)'} \qquad (t \text{ umbral})$$

we have

$$\frac{\partial x^{(r)'}}{\partial y^{(t)'}} = \frac{\partial x^{(r)}}{\partial y^{(t)}}$$

so that

$$\frac{\partial F}{\partial y^{(s)'}} \equiv \frac{\partial F}{\partial x^{(r)'}}\frac{\partial x^{(r)'}}{\partial y^{(s)'}} \qquad (r \text{ umbral})$$

$$\equiv \frac{\partial F}{\partial x^{(r)'}}\frac{\partial x^{(r)}}{\partial y^{(s)}}$$

showing that $\dfrac{\partial F}{\partial x^{(r)'}} \equiv X_r$ is a covariant tensor of rank one.

Suppose we wish to find the *geodesics* of our metrical space S_n. These are the curves for which the first variation of the length integral is zero.

$$F \equiv \sqrt{g_{lm} x^{(l)'} x^{(m)'}}$$

the g_{lm} being functions of position. We shall find it convenient

to take as parameter u the arc distance s along the sought-for geodesic.* Then when we put $\alpha = 0$ after the differentiations $F = 1$, from the definition of arc distance s, and we have

$$\left(\frac{\partial F}{\partial x^{(t)}}\right)_{\alpha=0} = \frac{1}{2}\frac{\partial g_{lm}}{\partial x^{(t)}}\dot{x}^{(l)}\dot{x}^{(m)}$$

where we use $\dot{x}^{(l)}$ to denote $\left(\frac{dx^{(l)}}{ds}\right)_{\alpha=0}$ so that $X^{(r)} \equiv (\dot{x}^{(r)})$ is the unit contravariant direction tensor along the sought-for geodesic.

Also

$$\left(\frac{\partial F}{\partial x^{(t)\prime}}\right)_{\alpha=0} = \frac{1}{2}\cdot 2g_{tm}(\dot{x}^{(m)})$$

and our equations are

$$\begin{aligned}\frac{1}{2}\frac{\partial g_{lm}}{\partial x^{(t)}}\dot{x}^{(l)}\dot{x}^{(m)} &= \frac{d}{ds}\cdot(g_{tm}\dot{x}^{(m)})\\ &= g_{tm}\ddot{x}^{(m)} + \dot{x}^{m}\frac{\partial g_{tm}}{\partial x^{(r)}}\dot{x}^{r}\end{aligned}$$

or

$$g_{tm}\ddot{x}^{m} + \dot{x}^{(r)}\dot{x}^{(m)}\left(\frac{\partial g_{tm}}{\partial x^{(r)}} - \frac{1}{2}\frac{\partial g_{rm}}{\partial x^{(t)}}\right) \equiv 0$$

$$(r,\ m \text{ umbral};\ t = 1 \cdots n)$$

Multiply through by g^{pt} and use t as an umbral symbol so as to obtain the n components of a contravariant tensor of rank one

$$\ddot{x}^{p} + g^{pt}\dot{x}^{(r)}\dot{x}^{m}\left(\frac{\partial g_{tm}}{\partial x^{(r)}} - \frac{1}{2}\frac{\partial g_{rm}}{\partial x^{(t)}}\right) = 0 \quad (r,\ m,\ t \text{ umbral})$$

It is now convenient to introduce the Christoffel three-index symbols of the first and second kinds defined as follows:

$$(a) \qquad [rs,\ t] \equiv [sr,\ t] \equiv \frac{1}{2}\left[\frac{\partial g_{rt}}{\partial x^{(s)}} + \frac{\partial g_{st}}{\partial x^{(r)}} - \frac{\partial g_{rs}}{\partial x^{(t)}}\right]$$

$$(b) \qquad \{rs,\ t\} \equiv \{sr,\ t\} \equiv g^{tp}[rs,\ p]$$

* However, this rules out those *minimal* geodesics along which s is constant.

which equations imply

$$g_{tq}\{rs, t\} \equiv g_{tq}g^{tp}[rs, p] \qquad (t, p \text{ umbral})$$
$$\equiv [rs, q]$$

Equations (a) give

$$[rs, t] + [rt, s] = \frac{\partial g_{st}}{\partial x^{(r)}}$$

Then we may write

$$\dot{x}^{(r)}\dot{x}^{m}\left(\frac{\partial g_{tm}}{\partial x^{(r)}} - \frac{1}{2}\frac{\partial g_{rm}}{\partial x^{(t)}}\right) \qquad (r, m \text{ umbral})$$

$$\begin{aligned}
&\equiv \dot{x}^{(r)}\dot{x}^{(m)}[[rt, m] + [rm, t] - \tfrac{1}{2}[rt, m] - \tfrac{1}{2}[tm, r]] \\
&\equiv \dot{x}^{(r)}\dot{x}^{(m)}[\tfrac{1}{2}[rt, m] + [rm, t] - \tfrac{1}{2}[mt, r]] \\
&\equiv \dot{x}^{(r)}\dot{x}^{(m)}[rm, t]
\end{aligned}$$

since an interchange of the umbral symbols r, m in the last three-index symbol leaves the summation unaltered.

Accordingly, on using the definition (b), the differential equations of the geodesics are

$$\ddot{x}^{(p)} + \{rm, p\}\dot{x}^{r}\dot{x}^{m} = 0 \qquad (p = 1 \cdots n)$$

From their derivation we know that these equations are contravariant of rank one. We proceed now to obtain a general rule which makes the tensor character of equations of this type apparent on inspection.

2. The Formulæ for Covariant Differentiation

From the covariant character of the g_{rs} we have

$$f_{rs} \equiv g_{lm}\frac{\partial x^{(l)}}{\partial y^{(r)}}\frac{\partial x^{(m)}}{\partial y^{(s)}} \qquad (l, m \text{ umbral})$$

$$\therefore \quad \frac{\partial f_{rs}}{\partial y^{(t)}} \equiv \frac{\partial g_{lm}}{\partial x^{(n)}}\frac{\partial x^{(l)}}{\partial y^{(r)}}\frac{\partial x^{(m)}}{\partial y^{(s)}}\frac{\partial x^{(n)}}{\partial y^{(t)}} + g_{lm}\left(\frac{\partial^2 x^{(l)}}{\partial y^{(r)}\partial y^{(t)}}\frac{\partial x^{(m)}}{\partial y^{(s)}} + \frac{\partial x^{(l)}}{\partial y^{(r)}}\frac{\partial^2 x^{(m)}}{\partial y^{(s)}\partial y^{(t)}}\right) \qquad (l, m, n \text{ umbral})$$

where in the differentiation we have remembered that g_{lm} is a

function of the y's only indirectly through the x's. We easily obtain two other similar equations by merely interchanging (r, t) and (s, t) in turn. We are careful to so distribute the umbral symbols l, m, n as to facilitate combination of the three equations obtained in this way. Thus

$$\frac{\partial f_{ts}}{\partial y^{(r)}} \equiv \frac{\partial g_{nm}}{\partial x^{(l)}} \frac{\partial x^{(l)}}{\partial y^{(r)}} \frac{\partial x^{(m)}}{\partial y^{(s)}} \frac{\partial x^{(n)}}{\partial y^{(t)}} + g_{lm}\left(\frac{\partial^2 x^{(l)}}{\partial y^{(r)}\partial y^{(t)}} \frac{\partial x^{(m)}}{\partial y^{(s)}} + \frac{\partial^2 x^{(m)}}{\partial y^{(s)}\partial y^{(r)}} \frac{\partial x^{(l)}}{\partial y^{(t)}}\right)$$

Now adding the first two of the equations and subtracting the third we have, on writing

$$[rt, s]' \equiv \frac{1}{2}\left[\frac{\partial f_{rs}}{\partial y^{(t)}} + \frac{\partial f_{ts}}{\partial y^{(r)}} - \frac{1}{2}\frac{\partial f_{rt}}{\partial y^{(s)}}\right], \quad \text{etc.}$$

$$[rt, s]' = [ln, m]\frac{\partial x^{(l)}}{\partial y^{(r)}} \frac{\partial x^{(m)}}{\partial y^{(s)}} \frac{\partial x^{(n)}}{\partial y^{(t)}} + g_{lm}\frac{\partial^2 x^{(l)}}{\partial y^{(r)}\partial y^{(t)}} \frac{\partial x^{(m)}}{\partial y^{(s)}}$$

[l, m, n umbral]

Now

$$g_{lm} = f_{pq}\frac{\partial y^{(p)}}{\partial x^{(l)}} \frac{\partial y^{(q)}}{\partial x^{(m)}}$$

from its covariant character (p, q umbral)

$$\therefore \quad g_{lm}\frac{\partial x^{(m)}}{\partial y^{(s)}} = f_{ps}\frac{\partial y^{(p)}}{\partial x^{(l)}}$$

To remove the coefficient of $\dfrac{\partial^2 x^{(l)}}{\partial y^{(r)}\partial y^{(t)}}$ multiply across by $f^{sk}\dfrac{\partial x^{(j)}}{\partial y^{(k)}}$ and make s and k umbral when we get

$$\{rt, k\}'\frac{\partial x^{(j)}}{\partial y^{(k)}} = g^{mj}[ln, m]\frac{\partial x^{(l)}}{\partial y^{(r)}} \frac{\partial x^{(n)}}{\partial y^{(t)}} + \frac{\partial^2 x^{(j)}}{\partial y^{(r)}\partial y^{(t)}}$$

from the relation (contravariant)

$$g^{mj} = f^{sk}\frac{\partial x^{(m)}}{\partial y^{(s)}} \frac{\partial x^{(j)}}{\partial y^{(k)}} \qquad (s,\ k \text{ umbral})$$

Finally

$$\frac{\partial^2 x^{(j)}}{\partial y^{(r)}\partial y^{(t)}} = \{rt, k\}' \frac{\partial x^{(j)}}{\partial y^{(k)}} - \{ln, j\} \frac{\partial x^{(l)}}{\partial y^{(r)}} \frac{\partial x^{(n)}}{\partial y^{(t)}}$$

from which on interchanging the rôle of the x and y coordinates we have

$$\frac{\partial^2 y^{(j)}}{\partial x^{(r)}\partial x^{(t)}} = \{rt, k\} \frac{\partial y^{(j)}}{\partial x^{(k)}} - \{ln, j\}' \frac{\partial y^{(l)}}{\partial x^{(r)}} \frac{\partial y^{(n)}}{\partial x^{(t)}}$$

Suppose now we have a covariant tensor of rank one X_r so that

$$Y_s \equiv X_r \frac{\partial x^{(r)}}{\partial y^{(s)}} \qquad (r \text{ umbral})$$

Then

$$\frac{\partial Y_s}{\partial y^{(t)}} \equiv X_r \frac{\partial^2 x^{(r)}}{\partial y^{(s)}\partial y^{(t)}} + \frac{\partial X_r}{\partial x^{(p)}} \frac{\partial x^{(p)}}{\partial y^{(t)}} \frac{\partial x^{(r)}}{\partial y^{(s)}}$$

$$\equiv X_r \left\{ \{st, k\}' \frac{\partial x^{(r)}}{\partial y^k} - \{lm, r\} \frac{\partial x^{(l)}}{\partial y^{(s)}} \frac{\partial x^{(m)}}{\partial y^{(t)}} \right\} + \frac{\partial X_r}{\partial x^{(p)}} \frac{\partial x^{(p)}}{\partial y^{(t)}} \frac{\partial x^{(r)}}{\partial y^{(s)}}$$

$$\therefore \quad \frac{\partial Y_s}{\partial y^{(t)}} - Y_k\{st, k\}' \equiv \left[\frac{\partial X_r}{\partial x^{(p)}} - X_k\{rp, k\}\right] \frac{\partial x^{(r)}}{\partial y^{(s)}} \frac{\partial x^{(p)}}{\partial y^{(t)}}$$

on altering suitably the umbral symbols lm to rp. These equations state that

$$\frac{\partial X_r}{\partial x^{(p)}} - X_k\{rp, k\} \equiv X_{rp}$$

is a covariant tensor of rank two. Consider now a contravariant tensor of rank one so that

$$Y^s \equiv X^r \frac{\partial y^{(s)}}{\partial x^{(r)}} \qquad (r \text{ umbral})$$

Then

$$\frac{\partial Y^{(s)}}{\partial y^{(t)}} \equiv \frac{\partial X^{(r)}}{\partial x^{(p)}} \frac{\partial x^{(p)}}{\partial y^{(t)}} \frac{\partial y^{(s)}}{\partial x^{(r)}} + X^r \frac{\partial^2 y^{(s)}}{\partial x^{(r)}\partial x^{(p)}} \frac{\partial x^{(p)}}{\partial y^{(t)}} \qquad (r, p \text{ umbral})$$

$$\equiv \frac{\partial X^{(r)}}{\partial x^{(p)}} \frac{\partial x^{(p)}}{\partial y^{(t)}} \frac{\partial y^{(s)}}{\partial x^{(r)}} + X^r \frac{\partial x^{(p)}}{\partial y^{(t)}} \left[\{rp, k\} \frac{\partial y^{(s)}}{\partial x^{(k)}} - \{lm, s\}' \frac{\partial y^{(l)}}{\partial x^{(r)}} \frac{\partial y^{(m)}}{\partial x^{(p)}} \right]$$

$$\therefore \frac{\partial Y^{(s)}}{\partial y^{(t)}} + Y^l\{lt, s\}' \equiv \left[\frac{\partial X^{(r)}}{\partial x^{(p)}} + X^k\{kp, r\}\right]\frac{\partial x^{(p)}}{\partial y^{(t)}}\frac{\partial y^{(s)}}{\partial x^{(r)}}$$

These equations state that

$$\frac{\partial X^{(r)}}{\partial x^{(p)}} + X^k\{kp, r\} \equiv X_p^{\ r}$$

is a mixed tensor of rank two.

These tensors of rank two are called the *covariant derivatives* of the covariant and contravariant tensors X_r and X^r respectively. Similar analysis can be carried out to obtain the covariant derivative of a tensor of any rank and character. To make this perfectly clear let us take the case of a mixed tensor $X_s^{\ r}$ of rank two:

$$Y_s^{\ r} \equiv X_q^{\ p}\frac{\partial x^{(q)}}{\partial y^{(s)}}\frac{\partial y^{(r)}}{\partial x^{(p)}}$$

$$\frac{\partial Y_s^{(r)}}{\partial y^{(t)}} \equiv \frac{\partial X_q^{\ p}}{\partial x^{(l)}}\frac{\partial x^{(l)}}{\partial y^{(t)}}\frac{\partial x^{(q)}}{\partial y^{(s)}}\frac{\partial y^{(r)}}{\partial x^{(p)}}$$
$$+ X_q^{\ p}\left[\frac{\partial^2 x^{(q)}}{\partial y^{(s)}\partial y^{(t)}}\frac{\partial y^{(r)}}{\partial x^{(p)}} + \frac{\partial x^{(q)}}{\partial y^{(s)}}\frac{\partial^2 y^{(r)}}{\partial x^{(p)}\partial x^{(l)}}\frac{\partial x^{(l)}}{\partial y^{(t)}}\right]$$
$$\equiv \frac{\partial X_q^{\ p}}{\partial x^{(l)}}\frac{\partial x^{(l)}}{\partial y^{(t)}}\frac{\partial x^{(q)}}{\partial y^{(s)}}\frac{\partial y^{(r)}}{\partial x^{(p)}}$$
$$+ X_q^{\ p}\left[\frac{\partial y^{(r)}}{\partial x^{(p)}}\left\{\{st, k\}'\frac{\partial x^{(q)}}{\partial y^{(k)}} - \{lm, q\}\frac{\partial x^{(l)}}{\partial y^{(s)}}\frac{\partial x^{(m)}}{\partial y^{(t)}}\right\}\right.$$
$$\left.+ \frac{\partial x^{(q)}}{\partial y^{(s)}}\frac{\partial x^{(l)}}{\partial y^{(t)}}\left\{\{pl, k\}\frac{\partial y^{(r)}}{\partial x^{(k)}} - \{nm, r\}'\frac{\partial y^{(n)}}{\partial x^{(p)}}\frac{\partial y^{(m)}}{\partial x^{(l)}}\right\}\right]$$

whence

$$\frac{\partial Y_s^{\ r}}{\partial y^{(t)}} - Y_k^{\ r}\{st, k\}' + Y_s^{\ k}\{kt, r\}'$$
$$\equiv \left\{\frac{\partial X_q^{\ p}}{\partial x^{(l)}} - X_k^{\ p}\{ql, k\} + X_q^{\ k}\{kl, p\}\right\}\frac{\partial y^{(r)}}{\partial x^{(p)}}\frac{\partial x^{(q)}}{\partial y^{(s)}}\frac{\partial x^{(l)}}{\partial y^{(t)}}$$

expressing that $\dfrac{\partial X_q^{\ p}}{\partial x^{(l)}} - X_k^{\ p}\{ql, k\} + X_q^{\ k}\{kl, p\}$ is a mixed tensor of rank three being covariant of rank two and contra-

variant of rank one. In general, the covariant derivative of $X^{r_1 \cdots r_m}_{s_1 \cdots s_n}$ is

$$X^{r_1 \cdots r_m}_{s_1 \cdots s_n s} \equiv \frac{\partial X^{r_1 \cdots r_m}_{s_1 \cdots s_n}}{\partial x^{(s)}} - X^{r_1 \cdots r_m}_{k s_2 \cdots s_n}\{s_1 s, k\} - \cdots - X^{r_1 \cdots r_m}_{s_1 \cdots s_{n-1} k}\{s_n s, k\}$$
$$+ X^{k r_2 \cdots r_m}_{s_1 \cdots s_n}\{ks, r_1\} + \cdots + X^{r_1 \cdots r_{m-1} k}_{s_1 \cdots s_n}\{ks, r_m\}$$

It will be noticed that + signs go with the contravariant symbols and negative with the covariant. Also the new label s is always second in the three-index symbols; the umbral label is first if taken from the contravariant and third if taken from the covariant indices.

3. Applications of the Rule of Covariant Differentiation

(a) *Riemann's four-index symbols and Einstein's Gravitational Tensor*

From any covariant tensor X_r we obtain as its covariant derivative

$$X_{rs} \equiv \frac{\partial X_r}{\partial x^{(s)}} - X_k\{rs, k\} \qquad (k \text{ umbral})$$

and as its second covariant derivative

$$X_{rst} \equiv \frac{\partial}{\partial x^{(t)}}\left[\frac{\partial X_r}{\partial x^{(s)}} - X_k\{rs, k\}\right] - X_{ps}\{rt, p\} - X_{rp}\{st, p\}$$
$$\equiv \frac{\partial^2 X_r}{\partial x^{(s)}\partial x^{(t)}} - X_k \frac{\partial}{\partial x^{(t)}}\{rs, k\} - \{rs, k\}\frac{\partial X_k}{\partial x^t}$$
$$- \frac{\partial X_p}{\partial x^{(s)}}\{rt, p\} + X_k\{ps, k\}\{rt, p\}$$
$$- \frac{\partial X_r}{\partial x^p}\{st, p\} + X_k\{rp, k\}\{st, p\}$$

From this by the elementary rule (*b*), Ch. 2, § 2, of tensor algebra we derive a new covariant tensor $\overline{X}_{rst} \equiv X_{rts}$ and the difference of these is a covariant tensor of rank 3 by rule (*a*), Ch. 2, § 1; i.e.,

$$\overline{\overline{X}}_{rst} \equiv X_{rst} - X_{rts}$$

whence

$$\overline{\overline{X}}_{rst} \equiv X_k \left[\frac{\partial}{\partial x^{(s)}} \{rt, k\} - \frac{\partial}{\partial x^{(t)}} \{rs, k\} + \{ps, k\}\{rt, p\} - \{pt, k\}\{rs, p\} \right]$$

the terms involving the derivatives of the X_r cancelling completely out. Now X_k is an *arbitrary* covariant tensor of rank one and so by the rule (*e*), Ch. 2, § 5—the converse of the rule (*d*) of composition—

$$\frac{\partial}{\partial x^{(s)}} \{rt, k\} - \frac{\partial}{\partial x^{(t)}} \{rs, k\} + \{ps, k\}\{rt, p\} - \{pt, k\}\{rs, p\} \equiv X_{rst}{}^k$$

is a mixed tensor of rank four of the type indicated by the positions of the labels.

If we write $k = t$ and use t as an umbral symbol we derive by rule (*d*) *Einstein's gravitational covariant tensor of rank two*

$$G_{rs} \equiv \frac{\partial}{\partial x^{(s)}} \{rt, t\} - \frac{\partial}{\partial x^{(t)}} \{rs, t\} + \{ps, t\}\{rt, p\} - \{pt, t\}\{rs, p\}$$

The mixed tensor $X_{rst}{}^k$ is usually denoted by the symbol $\{rk, ts\}$ and is known as the Riemann four-index symbol of the second kind. From it we obtain by the rule of composition the covariant tensor of rank four

$$[rj, ts] \equiv g_{jk} X_{rst}{}^k \equiv g_{jk}\{rk, ts\} \qquad (k \text{ umbral})$$

which is known as the Riemann four-index symbol of the first kind. From Einstein's tensor of rank two we obtain the invariant

$$G \equiv g^{rs} G_{rs} \qquad (r, s \text{ umbral})$$

which has been called the *Gaussian* or *total curvature of the space.* This name is given since G is regarded as a generalization of the expression given by Gauss for the curvature of a surface (i.e.,

$n = 2$). The term curvature is widely used in the literature of Relativity and so it may be well, in order to avoid a possible confusion of ideas on the subject, to discuss briefly what is meant by the curvature of a metrical space. To do this it is necessary to say a few words about the four-index symbols.

We have, by definition,

$$[pq, rs] \equiv g_{qk}\{pk, rs\} \equiv g_{qk}\left[\frac{\partial}{\partial x^{(s)}}\{pr, k\} - \frac{\partial}{\partial x^{(r)}}\{ps, k\} + \{pr, t\}\{ts, k\} - \{ps, t\}\{tr, k\}\right]$$

Recalling that

$$[pr, q] = g_{qk}\{pr, k\}$$

we have

$$g_{qk}\frac{\partial}{\partial x^{(s)}}\{pr, k\} = \frac{\partial}{\partial x^{(s)}}[pr, q] - \{pr, k\}\frac{\partial g_{qk}}{\partial x^{(s)}} \quad (k \text{ umbral})$$

$$= \frac{\partial}{\partial x^{(s)}}[pr, q] - \{pr, k\}([qs, k] + [ks, q])$$

from definition of $[qs, k]$ so that on operating similarly with $g_{qk}\dfrac{\partial}{\partial x^{(r)}}\{ps, k\}$ and writing $g_{qk}\{ts, k\} \equiv [ts, q]$ we find

$$[pq, rs] \equiv \frac{\partial}{\partial x^{(s)}}[pr, q] - \frac{\partial}{\partial x^{(r)}}[ps, q] - \{pr, k\}[qs, k] + \{ps, k\}[qr, k]$$

(the terms $\{pr, t\}[ts, q]$ and $-\{pr, k\}[ks, q]$ cancel since t and k are merely umbral symbols). Finally, in terms of the three-index symbols of the first kind,

$$[pq, rs] \equiv \frac{\partial}{\partial x^{(s)}}[pr, q] - \frac{\partial}{\partial x^{(r)}}[ps, q] + g^{kj}([ps, j][qr, k] - [pr, j][qs, k]) \quad (k, j \text{ umbral})$$

Writing out, in the first two terms of this expression, the values of the symbols, e.g.,

$$[pr, q] = \frac{1}{2}\left[\frac{\partial}{\partial x^{(p)}}g_{rq} + \frac{\partial}{\partial x^{(r)}}g_{pq} - \frac{\partial}{\partial x^{(q)}}g_{pr}\right]$$

we find

$$[pq, rs] \equiv \frac{1}{2}\left[\frac{\partial^2}{\partial x^{(p)}\partial x^{(s)}} g_{qr} - \frac{\partial^2}{\partial x^{(q)}\partial x^{(s)}} g_{pr} + \frac{\partial^2}{\partial x^{(q)}\partial x^{(r)}} g_{ps} \right.$$
$$\left. - \frac{\partial^2}{\partial x^{(p)}\partial x^{(r)}} g_{qs}\right] + g^{kj}([ps, j][qr, k] - [pr, j][qs, k])$$

From this formula it is apparent that

(*a*) An interchange of the indices or labels p, q merely changes the sign of the symbol.

$$[pq, rs] + [qp, rs] \equiv 0$$

(*b*) Similarly

$$[pq, rs] + [pq, sr] \equiv 0$$

(*c*) A complete reversal of the order of the labels does not alter the symbol $[pq, rs] \equiv [sr, pq]$. This depends on the symmetry relations $g^{kj} \equiv g^{jk}$.

(*d*) If we keep the first label fixed and permute the other three *cyclically* we get 3 symbols whose sum is identically zero, i. e.,

$$[pq, rs] + [pr, sq] + [ps, qr] \equiv 0$$

The number of non-vanishing symbols which are linearly distinct now follows. If $p = q$ or $r = s$ the symbol vanishes on account of (*a*) and (*b*). The number of choices for the first pair (p, q) is $n_2 = \dfrac{n(n-1)}{2}$ and similarly for the second pair (r, s). However relation (*c*) shows us that we do not get n_2^2 symbols by combining the two choices but

$$n_2^2 - \tfrac{1}{2}n_2(n_2 - 1) = \tfrac{1}{2}n_2(n_2 + 1)$$

The relation (*d*) will still further reduce the number of linearly distinct symbols. When the indices or labels p, q, r, s have numerical values which are not all distinct the relation (*d*) merely reduces to a combination of the relations (*a*), (*b*), (*c*). There are therefore $n(n-1)(n-2)(n-3)$ new relations in (*d*). However since there are three letters q, r, s permuted cyclically, each

relation will occur three times. Each of the relations (*a*), (*b*), (*c*) reduces the number $\dfrac{n(n-1)(n-2)(n-3)}{3}$ which remains in half and so there are

$$n_4 \equiv n(n-1)(n-2)(n-3) \div 24 \text{ distinct relations } (d).$$

There are accordingly but

$$\tfrac{1}{2}(n_2)(n_2+1) - n_4 = \frac{n^2(n^2-1)}{12}$$

distinct Riemann four-index symbols. For $n = 2$ there is but one which we may write [12, 12]. When we change the coordinates from x to y we have

$$[12, 12]_y \equiv [pq, rs]\frac{\partial x^{(p)}}{\partial y^{(1)}}\frac{\partial x^{(q)}}{\partial y^{(2)}}\frac{\partial x^{(r)}}{\partial y^{(1)}}\frac{\partial x^{(s)}}{\partial y^{(2)}}$$

(from covariant character)

Since there is but one distinct symbol $[pq, rs]$ it will factor out on the right and we get (since there are but four of the symbols which do not vanish)

$$[12, 12]_y = [12, 12]\cdot J^2$$

where J is the Jacobian $\dfrac{\partial(x^{(1)}x^{(2)})}{\partial(y^{(1)}y^{(2)})}\cdot$ We have already seen that $f = gJ^2$ and on division we obtain the *invariant* $K \equiv \dfrac{[12, 12]}{g}\cdot$

It is this invariant which Gauss called the total curvature of the space of *two* dimensions under discussion.

In order to compare this with the invariant

$$g^{rs}G_{rs} \qquad (r, s \text{ umbral};\ n = 2)$$

we have

$$G_{rs} \equiv \{rt, ts\} \qquad (t \text{ umbral})$$

$$\equiv g^{tp}[rpts]$$

$$\therefore\ G_{11} = g^{22}[1221]$$

$\{$since if $p = 1$ or $t = 1$, $[1pt1] \equiv 0$ by relations (*a*) and (*b*),$\}$

$$= -\, g_{11}[12, 12] \div g$$

from definition of g^{22},

$$= - g_{11}\cdot K$$

Similarly

$$G_{12} = g^{12}[12, 12] = - g_{21}\cdot K = - g_{12}\cdot K$$
$$G_{21} = g^{21}[21, 21] = - g_{12}\cdot K = - g_{21}\cdot K$$

from relation (c),

$$G_{22} = g^{11}[21, 12] = - g_{22}\cdot K$$

so that

$$g^{rs}G_{rs} = - Kg^{rs}g_{rs} = - 2K$$

since

$$g^{rs}G_{rs} = - K\{g^{r1}g_{r1} + g^{r2}g_{r2}\} = - 2K \quad (r \text{ umbral})$$

For a space in which, in some particular coordinate system x, the coefficients g_{rs} are constants all the three-index symbols $[pr, s]$ and in consequence all the symbols $\{pr, s\}$ and also the four-index symbols $[pq, rs]$ and $\{pq, rs\} = 0$. On account of the tensor character of these latter symbols we know that the Riemann tensors $[pq, rs]$ $\{pq, rs\}$ will be zero no matter what the coordinates are. Conversely the vanishing of the tensor $[pq, rs]$ expresses the fact that it is possible to find coordinates y such that the f_{rs} defined by the equations

$$f_{rs} \equiv g_{lm}\frac{\partial x^{(l)}}{\partial y^{(r)}}\frac{\partial x^{(m)}}{\partial y^{(s)}} \qquad (l, m \text{ umbral})$$

shall be constants. We may now apply the well-known method of reduction of a quadratic expression to a sum of squares (as in the determination of normal vibrations in dynamics where the expression for the kinetic energy is reduced to a sum of square terms); the transformations on the y's are linear in this operation and we finally get

$$(ds)^2 = \sum_{r=1}^{n} (dy^{(r)})^2$$

(If we restrict ourselves to real transformations there may be some negative squares; thus in the relativity theory there are three $-$ and one $+$ term.) A space of this character is said to be *Euclidean*

and the y's are called orthogonal Cartesian coordinates. Riemann defines curvature by means of his tensor $[pq, rs]$. When this tensor vanishes the curvature of the space is said to be zero so that *Euclidean space is one of Zero Riemann Curvature and conversely*. If the ratio of the component $[pq, rs]$ of the curvature tensor to the two-rowed determinant $\begin{vmatrix} g_{pr} & g_{ps} \\ g_{qr} & g_{qs} \end{vmatrix}$ is the same for all values of p, q, r, s, Riemann says the space is of constant curvature; otherwise the curvature will be different for different orientations at a point:* Gauss' total curvature, on the other hand, has a numerical value at each point in space and has nothing to do with the different orientations at that point. We may sum up by saying that a gravitational space is, at points free from matter, non-Euclidean, i.e., it has a Riemann curvature but its Gaussian curvature is *zero*.

It may be well to call attention to the fact that the definition

* The differential equations of the non-minimal geodesics of any space are

$$\frac{d^2x^{(r)}}{ds^2} + \begin{Bmatrix} lm \\ r \end{Bmatrix} \frac{dx^{(l)}}{ds}\frac{dx^{(m)}}{ds} = 0 \qquad (r = 1, \cdots, n;\ l, m \text{ umbral})$$

s being the arc length along the geodesic. It is known that the solutions $x^{(r)}$ of these equations are completely determined by the values of $x^{(r)}$ and $\frac{dx^{(r)}}{ds}$ for a particular value of s, $s = 0$ let us say. This is stated geometrically by saying that through any point in space there passes a unique geodesic with a given direction. If, now, through a definite point we construct the geodesics with the distinct directions $\xi^{(r)}$ and $\eta^{(r)}$ respectively $(r = 1, \cdots, n)$ and consider the family of geodesics through the point in question obtained by assigning to each a direction tensor whose rth component is proportional to $\lambda\xi^{(r)} + \mu\eta^{(r)}$ and then letting the ratio $\lambda : \mu$ vary, we obtain a *geodesic* spread V_2 of two dimensions which at the point in question has the orientation determined by the two directions ξ and η through the point. It is the curvature of this geodesic V_2 that Riemann calls the curvature of the space *relative* to the orientation determined by ξ and η. There is a remarkable theorem due to Schur (Math. Annalen, Bd. 27, p. 563, 1886) which says that if at every point the Riemann curvature of space is independent of the orientation the curvature at all points is the same. Such a space is, then, properly called a space of constant curvature.

Euclidean space given above is a "differential" definition; aces which are Euclidean according to this definition do not cessarily satisfy the postulate that one can proceed indefinitely a given direction without coming back to the starting point. he simplest example is the well-known one of a cylinder of unit dius. In this case $n = 2$, $y^{(1)} \equiv \phi$, the longitudinal angle easured in radians, and $y^{(2)} \equiv z$, the distance measured parallel the axes of the cylinder:

$$(ds)^2 \equiv (d\phi)^2 + (dz)^2 \equiv (dy^{(1)})^2 + (dy^{(2)})^2$$

the cylinder is cut along a generator and developed on a plane will cover a strip of breadth 2π on the plane. If we take ctangular Cartesian axes in the plane, with the $x^{(1)}$ axis parallel the strip, points whose $x^{(2)} \equiv \phi$ differ by 2π correspond to a nique point in the strip (that one with the same $x^{(1)}$) and to a nique point on the cylinder. Hence there are an infinity of traight lines (i.e., geodesics) joining any two points (with ifferent z's) on the cylinder. They develop into the ∞^1 straight nes joining the points

$$(x^{(1)}, x^{(2)}) \text{ and } (x^{(1)}, x^{(2)} + 2n\pi) \qquad (n = \pm 1, 2, \cdots)$$

n the plane. It is evident that speculations as to the "finiteess" or "infiniteness" of a space based on its *differential* haracteristics must be regarded with distrust.

CHAPTER VII

1. In this final chapter we shall treat in a brief way, as an application of the preceding analysis, the classical problems of Relativity. As in other applications of the methods of mathematical analysis to problems in physics the first, and here the most serious, difficulty is that of giving a physical significance to the coordinates. All systems of coordinates are, without doubt, equally valid for the statement of the laws of physics but not all are equally convenient. It is reasonable to suppose that for a given observer of phenomena a certain coordinate system may have a direct and simple relationship to the measurements he makes; such a coordinate system is called a natural system for that observer. It is necessary to define this natural system and to find by experience, or otherwise,* how the natural systems of different observers are related. This has been well done in the special or "Restricted Relativity Theory" but in the more general theory, which we propose to discuss here, much remains to be done in this part of the subject. In what follows we shall consider (*a*) the problem of determining the metrical character of the space-time continuum round a single gravitating center and (*b*) in consequence of the results of (*a*) the nature of the paths of a material particle and of a light ray in a gravitational field. We shall, following Einstein, make the fundamental assumption that the space which has a physical meaning or reality, i.e., with reference to which the laws of physics must have the tensor form (cf. Ch. 2, § 1), is one of four dimensions (commonly referred to as the Space-Time continuum).

* The relationship between the different systems may be arrived at by making various hypotheses whose truth or falsity must then be tested in the light of experience.

2. The Metrical Space Attached to a Single Gravitating Center

We assume that for an observer attached to the gravitating center one of the four coordinates, $x^{(4)}$ say, of his natural system is such that the coefficients g_{14}, g_{24}, g_{34} of the quadratic differential form for $(ds)^2$ vanish identically whilst those remaining are independent of $x^{(4)}$; $x^{(4)}$ is said to be a time coordinate and the field is said to be statical. Accordingly

$$(ds)^2 \equiv g_{44}(dx^{(4)})^2 + g_{lm}dx^{(l)}dx^{(m)} \quad (l, m = 1, 2, 3 \text{ umbral})$$

Now in any space of three dimensions we can always find orthogonal coordinate systems; for, writing the metrical $(ds)^2$ in its reciprocal form $(ds)^2 \equiv f^{rs}\eta_r\eta_s$, we have merely three equations $f^{rs} \equiv 0$ $(r \neq s)$—or explicitly

$$g^{lm}\frac{\partial y^{(r)}}{\partial x^{(l)}}\frac{\partial y^{(s)}}{\partial x^{(m)}} = 0$$

—to determine the three unknown functions y of x so that the coordinate curves y may be orthogonal. There is no lack of generality, then, in writing $(ds)^2$ for the statical field in the orthogonal form

$$(ds)^2 \equiv g_1(dx^{(1)})^2 + g_2(dx^{(2)})^2 + g_3(dx^{(3)})^2 + g_4(dx^{(4)})^2$$

where we have dropped the double labeling as unnecessary. (In general it is impossible to find orthogonal coordinates in space of *four* dimensions since there are now *six* differential equations $f^{rs} = 0$ $(r \neq s)$ for the *four* unknown functions y and these equations are not always consistent.) We must now go through the details of evaluating Einstein's gravitational tensor (cf. Ch. 6, § 3) for an orthogonal space.

The relations $g_{rs} = 0$; $g^{rs} = 0$ if $r \neq s$ make matters comparatively simple. We shall use r, s, t to denote *distinct* numerical values of the labels. Then

$$[rs,\,t] \equiv \frac{1}{2}\left(\frac{\partial g_{rt}}{\partial x^{(s)}} + \frac{\partial g_{st}}{\partial x^{(r)}} - \frac{\partial g_{rs}}{\partial x^{(t)}}\right) \equiv 0 \quad \text{by definition}$$

$$\{rs,\,t\} \equiv g^{tk}[rs,\,k] \equiv g^{tt}[rs,\,t] \equiv 0 \qquad (k \text{ umbral})$$

$$\{rs,\,r\} \equiv \{sr,\,r\} \equiv g^{rk}[sr,\,k] \equiv g^{rr}[sr,\,r] = \frac{1}{2g_r}\frac{\partial g_r}{\partial x^{(s)}}$$

(k being the only umbral label here)

$$\{rr,\,r\} = \frac{1}{2g_r}\frac{\partial g_r}{\partial x^{(r)}}$$

similarly and

$$\{rr,\,s\} = -\frac{1}{2g_s}\frac{\partial g_r}{\partial x^{(s)}}$$

The Riemann four-index symbol of the second kind (cf. Ch. 6, § 3) is defined by

$$\{pq,\,rs\} \equiv \frac{\partial}{\partial x^{(s)}}\{pr,\,q\} - \frac{\partial}{\partial x^{(r)}}\{ps,\,q\} + \{pr,\,l\}\{ls,\,q\} - \{ps,\,l\}\{lr,\,q\} \quad (l \text{ umbral})$$

and those components vanish identically, for an orthogonal coordinate system, where the pq, rs are distinct; [$\{pr,\,l\}$ vanishes unless $l = p$ or r in which case $\{ls,\,q\}$ vanishes]. To evaluate the remaining symbols write $r = q$ without, for the present, using q as an umbral symbol

$$\{pq,\,qs\} \equiv \frac{1}{2}\frac{\partial}{\partial x^{(s)}}\left(\frac{1}{g_q}\frac{\partial g_q}{\partial x^{(p)}}\right) + \frac{1}{4g_q{}^2}\frac{\partial g_q}{\partial x^{(p)}}\frac{\partial g_q}{\partial x^{(s)}} - \frac{1}{4g_p g_q}\frac{\partial g_p}{\partial x^{(s)}}\frac{\partial g_q}{\partial x^{(p)}} - \frac{1}{4g_s g_q}\frac{\partial g_s}{\partial x^{(p)}}\frac{\partial g_q}{\partial x^{(s)}}$$

The formulæ from this on take a simpler form if we use the symbols H defined by $g_r \equiv H_r{}^2$; thus

$$\{pq,\,qs\} \equiv \frac{1}{H_q}\frac{\partial^2 H_q}{\partial x^{(p)}\partial x^{(s)}} - \frac{1}{H_p H_q}\frac{\partial H_p}{\partial x^{(s)}}\frac{\partial H_q}{\partial x^{(p)}} - \frac{1}{H_s H_q}\frac{\partial H_s}{\partial x^{(p)}}\frac{\partial H_q}{\partial x^{(s)}}$$

Similarly we find

$$\{pq, qp\} \equiv \frac{1}{H_q}\frac{\partial^2 H_q}{\partial x^{(p)2}} + \frac{H_p}{H_q{}^2}\frac{\partial^2 H_p}{\partial x^{(q)2}} - \frac{H_p}{H_q{}^3}\frac{\partial H_p}{\partial x^{(q)}}\frac{\partial H_q}{\partial x^{(q)}}$$
$$- \frac{1}{H_pH_q}\frac{\partial H_p}{\partial x^{(p)}}\frac{\partial H_q}{\partial x^{(p)}} + \frac{H_p}{H_q}\left\{\frac{1}{H_r{}^2}\frac{\partial H_p}{\partial x^{(r)}}\frac{\partial H_q}{\partial x^{(r)}} + \frac{1}{H_s{}^2}\frac{\partial H_p}{\partial x^{(s)}}\frac{\partial H_q}{\partial x^{(s)}}\right\}$$

where r and s are the two labels different from p and q. The components of the Einstein tensor are now found by summing with respect to q. It will be recalled that $[pp, rs] = 0$ (p, r, s any values distinct or not, cf. Ch. 6, § 3). Hence

$$\{pp, rs\} \equiv g^{pk}[pk, rs] \equiv g^{pp}[pp, rs] = 0 \quad (k \text{ umbral})$$

Similarly $\{pq, ss\} = 0$, so that in forming G_{12}, for example, we need merely write

$$G_{12} \equiv \{13, 32\} + \{14, 42\}$$

whilst

$$G_{11} \equiv \{12, 21\} + \{13, 31\} + \{14, 41\}$$

It will be observed that differentiation with respect to $x^{(p)}$ and $x^{(s)}$ occurs in every term of $\{pq, qs\}$ and so the absence of the time coordinate $x^{(4)}$ from the coefficients makes G_{14}, G_{24}, G_{34} all identically zero.

We shall now make the following *hypotheses of symmetry*—(*a*) we shall suppose that the coordinate lines $x^{(1)}$ are geodesics of the space (all passing through the gravitating center). The equations of the non-singular geodesics have been found to be (Ch. 6, § 1)

$$\ddot{x}^{(r)} + \{lm, r\}\dot{x}^{(l)}\dot{x}^{(m)} = 0 \quad (r = 1, \cdots, 4;\ l, m \text{ umbral})$$

where dots denote differentiations with respect to the arc distance which we take as our coordinate $x^{(1)}$. Writing

$$\dot{x}^{(2)} = 0 = \dot{x}^{(3)} = \dot{x}^{(4)}, \qquad \dot{x}_1 = 1$$

(since $x^{(2)}$, $x^{(3)}$, $x^{(4)}$ are constant along the coordinate lines $x^{(1)}$) we find $\{11, r\} = 0$ which—from the values given for this symbol —yields $g_1 =$ constant. The constant is in fact unity since, by hypothesis, $ds = dx^{(1)}$ along the curves $x^{(2)} =$ const., $x^{(3)} =$ const., $x^{(4)} =$ const. It is apparent that it is sufficient that g_1 be a function of $x^{(1)}$ alone for we may make a change of variable $x^{(1)} \equiv x^{(1)}(y^{(1)})$ leaving the other coordinates unaltered; the argument shows conversely that if g_1 is a function of $x^{(1)}$ alone the coordinate lines $x^{(1)}$ are geodesics, the arc length along them being given by $s = \int \sqrt{g_1}\, dx^{(1)}$.

(*b*) $x^{(2)}$ and $x^{(3)}$ are directional coordinates serving to locate a point on the geodesic surface $x^{(1)} =$ const., $x^{(4)} =$ constant. We shall suppose that the arc differential on this surface (which may conveniently be called a geodesic sphere) cannot involve the "longitude" coordinate $x^{(3)}$ nor can the arc differential along a given "meridian" $x^{(3)} =$ constant depend on the "latitude" coordinate $x^{(2)}$. Hence g_2 is a function of $x^{(1)}$ alone whilst g_3 is a function of x_1 and x_2 alone.

(*c*) g_4 does not involve the directional coordinates $x^{(2)}$ and $x^{(3)}$ and so is a function of $x^{(1)}$ alone.

Accordingly, then, $x^{(3)}$ does not appear in the expression for $(ds)^2$ and so, in addition to $G_{14} = 0$, $G_{24} = 0$, $G_{34} = 0$ we have $G_{13} = 0$, $G_{23} = 0$. We must write down the five equations $G_{12} = 0$, $G_{11} = 0$, $G_{22} = 0$, $G_{33} = 0$, $G_{44} = 0$. The fact that H_4 is a function of $x^{(1)}$ alone and $H_1 = 1$ ($x^{(1)}$ being the arc distance along the geodesic curves $x^{(1)}$) gives $\{14, 42\} \equiv 0$ and from $G_{12} = \{13, 32\} = 0$ we get

$$\frac{\partial^2 H_3}{\partial x^{(1)} \partial x^{(2)}} - \frac{1}{H_2} \frac{\partial H_2}{\partial x^{(1)}} \frac{\partial H_3}{\partial x^{(2)}} = 0$$

which gives, on integration with respect to $x^{(1)}$,

$$\frac{1}{H_2} \frac{\partial H_3}{\partial x^{(2)}} = \text{a function independent of } x_1 \qquad (A)$$

$G_{11} = 0$ yields

$$\frac{1}{H_2}\frac{\partial^2 H_2}{\partial x^{(1)^2}} + \frac{1}{H_3}\frac{\partial^2 H_3}{\partial x^{(1)^2}} + \frac{1}{H_4}\frac{\partial^2 H_4}{\partial x^{(1)^2}} = 0 \qquad (B)$$

$G_{44} = 0$ gives

$$H_4\frac{\partial^2 H_4}{\partial x^{(1)^2}} + H_4\frac{\partial H_4}{\partial x^{(1)}}\left\{\frac{1}{H_2}\frac{\partial H_2}{\partial x^{(1)}} + \frac{1}{H_3}\frac{\partial H_3}{\partial x^{(1)}}\right\} = 0$$

which on integration with respect to $x^{(1)}$ gives

$$H_2 H_3 \frac{\partial H_4}{\partial x^{(1)}} \text{ independent of } x^{(1)} \qquad (C)$$

Eliminating $\dfrac{\partial H_3}{\partial x^{(2)}}$ between (C) and (A) we get $H_2^2 \dfrac{\partial H_4}{\partial x^{(1)}}$ independent of $x^{(1)}$. Since it cannot involve any variable but $x^{(1)}$ we have

$$H_2^2 H_4' = a \text{ constant } \alpha, \text{ let us say}; \qquad (C')$$

primes denoting differentiations with respect to $x^{(1)}$.

Again from (A) $\dfrac{\partial H_3}{\partial x^{(2)}} = H_2 \times$ a function of $x^{(2)} = H_2 \dfrac{\partial \phi}{\partial x^{(2)}}$ say where ϕ is a function of $x^{(2)}$ alone. Then $H_3 = H_2\phi + f$ where f is a function of $x^{(1)}$ alone. Now (B) shows that $\dfrac{1}{H_3}\dfrac{\partial^2 H_3}{\partial x^{(1)^2}}$ is a function of $x^{(1)}$ alone so that its derivative with respect to $x^{(2)}$ vanishes. Evaluating this derivative we find

$$\frac{\partial \phi}{\partial x^{(2)}}\{fH_2'' - f''H_2\} = 0$$

We can now proceed in various ways; either make ϕ a constant or $fH_2' - f'H_2$ (of which the second factor is the derivative) a constant giving $f = \text{const.} \times H_2 \int \dfrac{dx^{(1)}}{H_2^{(2)}}$. *We choose the latter alternative and make the constant zero* so that $f = 0$ giving $H_3 = H_2\phi$ where ϕ is a function of $x^{(2)}$ alone.

φ is determined by means of the equation $G_{22} = 0$. This gives

$$H_2 H_2'' + \frac{1}{H_3} \frac{\partial^2 H_3}{\partial x^{(2)^2}} + \frac{H_2}{H_3} \frac{\partial H_3}{\partial x^{(1)}} H_2' + \frac{H_2}{H_4} H_2' H_4' = 0 \qquad (D)$$

On substituting $H_3 = H_2 \phi$ in (D) we find that $\dfrac{1}{\phi} \dfrac{d^2\phi}{dx^{(2)^2}}$ is equal to a function of $x^{(1)}$ alone; but from its form and the definition of ϕ it cannot involve $x^{(1)}$ and so must be a constant. This constant may, by a proper choice of unit for $x^{(2)}$, be put either 1 or zero. We choose the first alternative and find, by suitably choosing the origin of measurement for $x^{(2)}$, $\phi = \sin x^{(2)}$.

$G_{33} = 0$ gives

$$H_2 \frac{\partial^2 H_3}{\partial x^{(1)^2}} + \frac{H_3}{H_2^2} \frac{\partial^2 H_3}{\partial x^{(2)^2}} + H_2 \frac{\partial H_3}{\partial x^{(1)}} \left\{ \frac{1}{H_2} \frac{\partial H_2}{\partial x^{(1)}} + \frac{1}{H_4} \frac{\partial H_4}{\partial x^{(1)}} \right\} = 0 \quad (E)$$

and on substituting $\phi = \sin x^{(2)}$, $H_3 = H_2 \phi$, both (D) and (E) yield the same equation

$$H_2 H_2'' + H_2 H_2' \left\{ \frac{H_2'}{H_2} + \frac{H_4'}{H_4} \right\} = 1 \qquad (D')$$

(B) gives

$$2 \frac{H_2''}{H_2} + \frac{H_4''}{H_4} = 0 \qquad (B')$$

On differentiating (C') and eliminating $H_2 H_4''$ we find

$$H_2'' H_4 = H_2' H_4'$$

which gives on integration

$$H_2' = \beta H_4$$

where β is an arbitrary constant.

Eliminating H_4 between (C') and (C'') we have $H_2'' = \dfrac{\alpha\beta}{H_2^2}$ which on integration gives

$$(H_2')^2 = 2\left(\gamma - \frac{\alpha\beta}{H_2}\right)$$

where γ is an arbitrary constant.

Putting $H_4 = (H_2')/\beta$ in (D') we have

$$2H_2H_2'' + (H_2')^2 = 1$$

so that $1 = 2\gamma$ giving $\gamma = \frac{1}{2}$ and hence finally H_2 is determined by the differential equation

$$(H_2')^2 = 1 - \frac{2\alpha\beta}{H_2}$$

and then

$$(ds)^2 = (dx^{(1)})^2 + H_2^2\{(dx^{(2)})^2 + \sin^2 x^{(2)}(dx^{(3)})^2\} + \frac{1}{\beta^2}(H_2')^2(dx^{(4)})^2$$

It is usual to change the coordinate $x^{(1)}$, leaving the others unaltered. We write $x^{(1)} = x^{(1)}(y^{(1)})$ where $y^{(1)} = H_2$.

$$(dy^{(1)})^2 = (H_2')^2(dx^{(1)})^2 = \left(1 - \frac{2\alpha\beta}{y^{(1)}}\right)(dx^{(1)})^2$$

and we have

$$(ds)^2 = \left(1 - \frac{2\alpha\beta}{y^{(1)}}\right)^{-1}(dy^{(1)})^2 + y^{(1)2}\{(dx^{(2)})^2 + \sin^2 x^{(2)}(dx^{(3)})^2\} + \frac{1}{\beta^2}\left(1 - \frac{2\alpha\beta}{y^{(1)}}\right)(dx^{(4)})^2$$

This is the form chosen by Einstein (that it is only one of many is evident from its derivation). If $\alpha\beta = 0$ it reduces to the well-known Euclidean form where $y^{(1)} = r$, $x^{(2)} = \theta$, $x^{(3)} = \phi$ are space polar coordinates. It remains to attach some physical significance to the constant $\alpha\beta$ and to take up the problem (*b*) stated at the beginning of this chapter. In order to conform to the usual notation we write henceforth $y^{(1)} = ir$; $x^{(2)} = \theta$; $x^{(3)} = \phi$; $x^{(4)} = t$ where $i^2 = -1$.

Choosing the unit of $x^{(4)}$ or t so that $\beta^2 = +1$ and writing $i\alpha = -a$ we have

$$(ds)^2 = -\left\{\left(1 - \frac{2a}{r}\right)^{-1}(dr)^2 + r^2(d\theta)^2 + r^2\sin^2\theta(d\varphi)^2\right\} + \left(1 - \frac{2a}{r}\right)(dt)^2$$

3. Determination of the Path of a Freely Moving Particle

A physical *law of inertia* is postulated to the effect that a *freely moving material particle* in a gravitational field *will* follow the non-minimal *geodesic lines* of the *four-dimensional space time continuum* which, for the single gravitating center, has the metrical geometry characterized by the form given above by $(ds)^2$.

A second postulate is that *rays of light follow the minimal geodesics*—those for which $ds = 0$. In the *ordinary* Euclidean space these lines are imaginary, i.e., have points with imaginary coordinates but the occurrence of the negative signs in the expression for $(ds)^2$ gives real minimal lines in our problem. For example, the light rays directed towards or away from the center, those for which θ and ϕ are constant, are characterized by the equation

$$\left(1 - \frac{2a}{r}\right)^{-1}(dr)^2 - \left(1 - \frac{2a}{r}\right)(dt)^2 = 0 \quad \text{or} \quad \frac{dr}{dt} = \pm\left(1 - \frac{2a}{r}\right)$$

In order, then, to solve the problem of the free motion of a material particle we have merely to determine the non-minimal geodesics whose equations are

$$\ddot{x}^{(p)} + \{lm, p\}\dot{x}^{(l)}\dot{x}^{(m)} = 0 \qquad \text{(cf. Ch. 6. § 1)}$$

the dots denoting differentiations with respect to the arc length along the geodesic. For an orthogonal space of four dimensions these simplify to four equations of the type

$$\begin{aligned}\ddot{x}^{(1)} + \{11, 1\}\dot{x}^{(1)^2} + \{22, 1\}(\dot{x}^{(2)})^2 + \{33, 1\}(\dot{x}^{(3)})^2 \\ + \{44, 1\}(\dot{x}^{(4)})^2 + 2\{12, 1\}\dot{x}^{(1)}\dot{x}^{(2)} + 2\{13, 1\}\dot{x}^{(1)}\dot{x}^{(3)} \\ + 2\{14, 1\}\dot{x}^{(1)}\dot{x}^{(4)} = 0\end{aligned}$$

However we need use only three of these equations, replacing the fourth by $g_{rs}\dot{x}^{(r)}\dot{x}^{(s)} = 1$ which is easily seen to be a consequence of the differential equations

$$\ddot{x}^{(r)} + \{lm, r\}\dot{x}^{(l)}\dot{x}^{(m)} = 0$$

we multiply these by g_{rs} and use r as an umbral symbol to ain

$$g_{rs}\ddot{x}^{(r)} + [lm, s]\dot{x}^{(l)}\dot{x}^{(m)} = 0$$

l then avail ourselves of the definition of the symbols , s]. (Ch. 6, § 1.) $\frac{d}{ds}(g_{rs}\dot{x}_r\dot{x}_s)$ is found to be zero). In our

blem it is convenient to omit the first of the four equations, other three simplifying, on using the values for the three-ex symbols given (Ch. 7, § 2), to

$$\left.\begin{aligned} -\frac{1}{2g_2}(\dot\phi)^2\frac{\partial g_3}{\partial\theta} + \frac{1}{g_2}\frac{\partial g_2}{\partial r}\dot{r}\dot{\theta} &= 0 \quad (A)\\ +\frac{1}{g_3}\frac{\partial g_3}{\partial r}\dot{r}\dot{\varphi} + \frac{1}{g_3}\frac{\partial g_3}{\partial\theta}\dot{\theta}\dot{\varphi} &= 0 \quad (B)\\ \ddot{t} + \frac{1}{g_4}\frac{\partial g_4}{\partial r}\dot{r}\dot{t} &= 0 \quad (C)\end{aligned}\right\} \quad \begin{aligned} \text{where } g_1 &= -\left(1-\frac{2a}{r}\right)^{-1};\\ g_2 &= -r^2;\\ g_3 &= -r^2\sin^2\theta;\\ g_4 &= +\left(1-\frac{2a}{r}\right).\end{aligned}$$

these we have to add the first integral

$$g_1\dot{r}^2 + g_2\dot{\theta}^2 + g_3\dot{\varphi}^2 + g_4\dot{t}^2 = 1 \qquad (D)$$

uations (B) and (C) are immediately integrable giving

$$g_3\dot\phi = \text{constant} = -h \text{ say} \qquad (B')$$

l

$$g_4\dot{t} = \text{constant} = +C \text{ say} \qquad (C')$$

on substituting the values of g_3 and g_4

$$r^2\sin^2\theta\dot\varphi = h; \qquad \left(1-\frac{2a}{r}\right)\dot{t} = C$$

uation (A) may be written

$$\frac{d}{ds}(r^2\dot\theta) - r^2\sin\theta\cos\theta(\dot\phi)^2 = 0$$

e now proceed to eliminate the parameter s and find a relation inecting θ and ϕ. Assuming that $\dot\phi \neq 0$ (ϕ = constant is a ecial case which is susceptible to the analysis given below on a

mere interchange of θ and φ) we have

$$\dot{\theta} = \frac{d\theta}{d\phi} \cdot \dot{\phi}$$

so that

$$\frac{d}{ds}(r^2\dot{\theta}) = \left\{2r\frac{dr}{d\phi}\frac{d\theta}{d\phi} + r^2\frac{d^2\theta}{d\phi^2}\right\}(\dot{\phi})^2 + r^2\frac{d\theta}{d\phi} \cdot \ddot{\phi}$$

On substituting the value of $\ddot{\phi}$ from (B) we have

$$\frac{d}{ds}(r^2\dot{\theta}) = \{2rr'\theta' + r^2\theta''\}\dot{\phi}^2 - r^2\theta'\left\{\frac{2}{r}r'\dot{\phi}^2 + 2\cot\theta\theta'\dot{\phi}^2\right\}$$

where we denote differentiations with respect to the new independent variable ϕ by primes. Equating this to $r^2\sin\theta\cos\theta(\dot{\phi})^2$ and dividing out by $r^2\dot{\phi}^2$ we obtain

$$\theta'' - 2\cot\theta(\theta')^2 = \sin\theta\cos\theta$$

If now we choose our directional coordinate θ so that initially

$$\theta = \frac{\pi}{2}, \qquad \theta' = 0$$

We see that $\theta'' = 0$ and then on differentiating the above equation with respect to φ, $\theta''' = 0$ and so for all the other derivatives, i.e., θ is a constant as φ varies. Otherwise expressed the general integral of the equation for θ as a function of φ is found by writing $z = \cot\theta$ yielding $z'' + z = 0$ to be $\cot\theta = L\cos(\phi + M)$ where L and M are arbitrary constants. We choose our initial conditions as above so that $L = 0$ giving $\theta \equiv \frac{\pi}{2}\cdot$ Putting in this value for θ we find

$$r^2\dot{\phi} = h \tag{B'}$$

$$\left(1 - \frac{2a}{r}\right)\dot{t} = C \tag{C''}$$

and from (D)

$$\left(1 - \frac{2a}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\phi}^2 - \left(1 - \frac{2a}{r}\right)\dot{t}^2 = -1 \tag{D'}$$

Just as in the usual Newtonian treatment of planetary motion, it is convenient to write $r = 1/u$ and to again use φ as the independent variable. We have

$$\dot{r} = -\dot{u}/u^2 = -u'\dot{\phi}/u^2 = -hu' \qquad \text{(from } B'')$$

and then (D') yields, on making use of (C''),

$$(u')^2 + u^2 = 2au^3 + \frac{C^2 - 1}{h^2} + \frac{2au}{h^2} \qquad (E)$$

Now, in the Newtonian treatment, the equation giving the path of a particle under a central force is

$$u'' + u = F/h^2u^2$$

where F is the acceleration towards the center and $h \equiv r^2 \dfrac{d\phi}{dt}$ is the constant of areas. Instead of this we have on differentiating the equation (E) just obtained

$$u'' + u = 3au^2 + \frac{a}{h^2}$$

so that we may, in a general manner, express Einstein's modification of the Newtonian law of gravitation by saying that there is superimposed to the inverse square law attraction an inverse fourth power attraction, the relative strength of the attracting masses being as $1 : 3h^2$. It remains to determine, at any rate approximately, the nature and magnitude of the constants a, h and C which arose in the integration of our differential equations. For large values of r, and therefore small values of u, the Newtonian law is a first approximation and so neglecting the term in u^2 in the equation for u'', $a = F/u^2 = \mu m$; μ being the gravitational constant and m the mass of the sun. Hence if we choose our unit of mass so that $\mu = 1$, $a = m$, where now m is what is known as the *gravitational mass* of the attracting center (notice that we have identified, for small values of u, our r and ϕ with the usual polar coordinates of Euclidean geometry). The velocity of light directly towards the attracting center is $1 - \dfrac{2a}{r}$ and accordingly

our unit of time is such that for small values of u the velocity of light is unity; i.e., if the unit of length be 1 cm., the unit of time employed is $1/c$ seconds where $c = 3.10^{10}$. In the theory of relativity there is no absolute distinction between space and time and so we refer to our time unit as one centimeter (1 cm. being the distance traversed by light in one time unit). It is to be observed that in Newtonian mechanics gravitational mass m has dimensions L^3T^{-2} so that if L and T have the same dimensions $a = m$ has the dimensions of a length. The equation

$$u'' + u = \frac{m}{h^2}$$

of the Newtonian theory yields

$$u - \frac{m}{h^2} = P \cos(\phi - \phi_0)$$

where P and ϕ_0 are arbitrary constants of integration.

Comparing this with the polar equation of a conic

$lu = 1 + e \cos \phi$ (l = semi-latus rectum, e = eccentricity)

we have $\dfrac{h^2}{m} = l = A(1 - e^2)$ where A is the semi-major axis. If T is the period of revolution

$$h = \frac{2 \times \text{Area of ellipse}}{T} = 2\pi A^2(1 - e^2)^{1/2}/T$$

whence

$$m = h^2/A(1 - e^2) = 4\pi^2 A^3/T^2 = \omega^2 A^3$$

where ω is the angular velocity of the planet. This gives for the sun m = 1.47 kilometers or $1.47.10^5$ cms. For the planets then m/r is a small quantity of the order 10^{-8}. In order to determine the constant C we differentiate

$$u = \frac{m}{h^2}(1 + e \cos \phi)$$

and find

$$(u')^2 + u^2 = \frac{m^2}{h^4}(1 + 2e \cos \phi + e^2) = \frac{2m}{h^2}u - \frac{m^2}{h^4}(1 - e^2)$$

and comparing this with the equation (E) we have

$$1 - C^2 = \frac{m^2}{h^2}(1 - e^2) = m/A$$

It is to be observed that the values of m, C and h obtained in this way are found from the Newtonian theory and so are to be regarded as first approximations. In particular we have identified the h of (B') with $r^2\dfrac{d\varphi}{dt}$ so that we have written $\dfrac{d\phi}{ds} = \dfrac{d\varphi}{dt}$

Accurately

$$\frac{d\varphi}{ds} = \frac{d\varphi}{dt}\cdot\frac{dt}{ds} = \frac{d\varphi}{dt}\cdot C\cdot\left(1 - \frac{2m}{r}\right)^{-1} \qquad \text{(from } C''\text{)}$$

But

$$C = \left(1 - \frac{m}{A}\right)^{1/2}$$

so that neglecting quantities of the order 10^{-8}

$$\frac{d\varphi}{ds} = \frac{d\varphi}{dt}$$

Substituting the expressions just obtained in (E) we have to integrate the first order differential equation

$$\left(\frac{du}{d\phi}\right)^2 = 2mu^3 - u^2 + 2mu/h^2 - m^2(1 - e^2)/h^4$$

This equation defines u as an elliptic function of φ; or inversely φ as an elliptic integral. It simplifies the algebra somewhat to write $mu = v$ and to put $m^2/h^2 = \alpha$. We have already seen that $m^2/h^2 = \dfrac{m}{A(1 - e^2)}$ so that if e is not very nearly equal to unity α is a small quantity of the same order of magnitude as m/A or 10^{-8}. Our equation is now

$$\left(\frac{dv}{d\varphi}\right)^2 = 2v^3 - v^2 + 2\alpha v - \alpha^2(1 - e^2)$$

Now the discriminant of the literal cubic

$$a_0v^3 + a_1v^2 + a_2v + a_3 = 0$$

is

$$a_1^2a_2^2 + 18a_0a_1a_2a_3 - 4a_0a_2^3 - 4a_1^3a_3 - 27a_0^2a_3^2$$

For the cubic on the right-hand side of the equation giving $(dv/d\varphi)^2$ this is

$$4\alpha^2e^2 + 8\alpha^3(1 - 9e^2) - 108\alpha^4(1 - e^2)^2.$$

On account of the small magnitude of α this is positive, the first term being the dominant one. Hence the cubic has three real roots which we denote, in descending order of magnitude, by v_1, v_2, v_3. When $\alpha = 0$ the roots are $\frac{1}{2}$, 0, 0, and so we try first $v = k\alpha$ and find $k = (1 - e)$ or $(1 + e)$ and then secondly $v = \frac{1}{2} + k\alpha$ and find $k = -2$. Hence, to a first approximation, the three roots of the cubic giving $(dv/d\varphi)^2$ are $v_3 = \alpha(1 - e)$; $v_2 = \alpha(1 + e)$; $v_1 = \frac{1}{2} - 2\alpha$. Further since $(dv/d\varphi)^2$ cannot be negative in the problem v must lie between v_3 and v_2 or between v_1 and $+\infty$. As r does not tend to zero v does not tend to ∞ and hence v lies between v_3 and v_2. We have

$$\varphi = \int_{v_3}^{v} \frac{dv}{\sqrt{2(v - v_1)(v - v_2)(v - v_3)}}.$$

The variable v oscillates between the values v_3 and v_2; at these values $dv/d\varphi = 0$, so that v has an extreme value; as v passes through the value v_2 retracing its values both dv and the radical change signs so that φ steadily increases. The change in φ between two successive extreme values of v, i.e., between perihelion and aphelion of the planet, is

$$\Delta\varphi = \int_{v_3}^{v_2} \frac{dv}{\sqrt{2(v - v_1)(v - v_2)(v - v_3)}}$$

It is convenient to make a simple linear transformation of the variable of integration. Write $v = a + bz$ and determine the coefficients a and b of the transformation so that to the roots v_3 and v_2 of the cubic will correspond values 0 and 1 of z respectively. The values are $a = v_3$; $b = v_2 - v_3$ and then the third root v_1

goes over into $z = \frac{1}{k^2}$ where $k^2 = \frac{v_2 - v_3}{v_1 - v_3}$. The cubic $2(v - v_1)$ $(v - v_2)(v - v_3)$ transforms into $2b^3z(1 - z)\left(\frac{1}{k^2} - z\right)$ so that

$$\Delta\varphi = \frac{k}{\sqrt{2b}}\int_0^1 \frac{dz}{\sqrt{z(1-z)(1-k^2z)}}$$

This simplifies considerably on writing $z = \sin^2\theta$ when in fact

$$\Delta\varphi = \frac{2k}{\sqrt{2b}}\int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}}$$

Now

$$k^2 = \frac{v_2 - v_3}{v_1 - v_3} = \frac{2\alpha e}{\frac{1}{2} - 3\alpha + e\alpha}$$

(to a first approximation) is a small quantity of the same order of magnitude as α; hence we can expand $(1 - k^2\sin^2\theta)^{-1/2}$ in a rapidly convergent series and a mere integration of the initial terms will give a very good approximation to $\Delta\varphi$. The multiplier of the integral is

$$2\sqrt{\frac{k^2}{2b}} = \sqrt{2/v_1 - v_3} = 2[1 - 2\alpha(3 - e)]^{-1/2} = 2[1 + \alpha(3 - e)]$$

and using $\int_0^{\pi/2}\sin^2\theta\, d\theta = \pi/4$ we find

$$\Delta\varphi = 2\{1 + \alpha(3 - e)\}\cdot\left\{\frac{\pi}{2} + \frac{k^2}{8}\pi\right\}$$

but $k^2 = 4\alpha e$ to a first approximation so

$$\Delta\varphi = \pi\{1 + \alpha(3 - e)\}\{1 + \alpha e\} = \pi\{1 + 3\alpha\}$$

Hence in a complete revolution the perihelion advances by an amount equal to

$$3\alpha = 3\frac{m^2}{h^2} = \frac{3m}{A(1 - e^2)} = \frac{12\pi^2A^2}{T^2(1 - e^2)}$$

of a complete revolution, T being the period in our units. If we

wish to use the period in seconds and measure A in kilometers then the unit of time in the formula given is the time it takes light to travel 1 kilometer $= 1/3.10^5$ seconds; hence if T is the period in seconds the fractional advance of the perihelion per revolution is $\dfrac{12\pi^2 A^2}{9.10^{10} T^2(1 - e^2)}$. On substituting the values of A, T, and e for Mercury's path this works out to be an advance of 43″ per century. For the other planets e is much smaller than for Mercury and the amount of advance of perihelion is much smaller; save in the case of Mars the predicted advance is too small to be detected by observation.

4. The Path of a Light Ray in the Gravitational Field of a Single Attracting Center

These paths satisfy the equation $(ds)^2 = 0$ or $ds = 0$; they are geodesics since, ds being the non-negative root of the expression for $(ds)^2$, no curve can have a negative length. The method of the preceding paragraph does not, however, immediately apply since the arc length s along a light ray, being a constant, cannot be used as an independent variable or parameter in terms of which the coordinates x may be expressed. Further in the discussion of Ch. 6, § 1, it was assumed that the integral

$$I(\alpha) \equiv \int_{C_\alpha} ds$$

could be expanded in a Taylor series in powers of α so that the existence of the derivative $(\partial I/\partial \alpha)_{\alpha=0}$ was presupposed. It is apparent, however, on differentiation of

$$ds = \sqrt{g_{lm} dx^{(l)} dx^{(m)}} \qquad (l,\ m \text{ umbral})$$

that if $ds = 0$ when $\alpha = 0$, $\dfrac{\partial}{\partial \alpha}(ds)$ becomes meaningless when $\alpha = 0$ on account of the zero factor $(ds)_{\alpha=0}$ which occurs in the denominator. These difficulties are overcome in the following manner. If we investigate those curves (non-minimal) for which

the first variation of the integral $I \equiv \int (ds)^2$* is zero we are led to exactly the same differential equations as those of Ch. 6, § 1, which express the fact that the first variation of $I \equiv \int ds$ is zero. Accordingly we now derive the equations of the minimal geodesics from the fact that the first variation of $I = \int (ds)^2$ is zero, ds being zero along the curves. The coordinates x are supposed expressed in terms of any convenient parameter v and differentiations with respect to this parameter are denoted by primes. The Euler-Lagrangian equations are (cf. Ch. 6, § 1)

$$\frac{\partial F}{\partial x^{(r)}} - \frac{d}{dv}\left(\frac{\partial F}{\partial x^{(r)'}}\right) = 0 \qquad (r = 1, \cdots, 4)$$

where

$$F \equiv (ds)^2 \equiv g_{lm} x^{(l)'} x^{(m)'} \qquad (l,\ m \text{ umbral})$$

Hence

$$\begin{aligned}\frac{\partial g_{lm}}{\partial x^{(r)}} x^{(l)'} x^{(m)'} &= 2\frac{d}{dv}\left(g_{rl} x^{(l)'}\right) \qquad (l,\ m \text{ umbral})\\ &= 2\left\{g_{rl} x^{(l)''} + \frac{\partial g_{rl}}{\partial x^{(m)}} x^{(l)'} x^{(m)'}\right\}\\ &= \left\{2 g_{rl} x^{(l)''} + \left(\frac{\partial g_{rl}}{\partial x^{(m)}} + \frac{\partial g_{rm}}{\partial x^{(l)}}\right) x^{(l)'} x^{(m)'}\right\}\end{aligned}$$

or

$$g_{rl} x^{(l)''} + [lm,\ r] x^{(l)'} x^{(m)'} = 0 \qquad (l,\ m \text{ umbral})$$

Multiplying by g^{pr} and using r as an umbral symbol we obtain

$$x^{(p)''} + \{lm,\ p\} x^{(l)'} x^{(m)'} = 0 \quad (p = 1, 2, 3, 4)$$

which are exactly the equations of Ch. 6, § 1. The first integral of these equations which has already been mentioned may be very briefly obtained as follows. Since $F \equiv g_{lm} x^{(l)'} x^{(m)'}$ is

* Attention should, however, be called to the fact that this integral is not, properly speaking, a line integral at all; its value depends not only on the curve over which it is extended but on the particular parametric mode of representation chosen for this curve. In order that the value of the integral should not be dependent on the parametric representation the integrand should be positively *homogeneous* of degree *unity* in the derivatives x'.

homogeneous of degree 2 in the x' we have, by Euler's theorem on homogeneous functions,

$$x^{(r)'}\frac{\partial F}{\partial x^{(r)'}} \equiv 2F$$

a result immediately verifiable directly (r umbral). On multiplying the equations

$$\frac{\partial F}{\partial x^{(r)}} - \frac{d}{dv}\left(\frac{\partial F}{\partial x^{(r)'}}\right) = 0$$

by $x^{(r)'}$ and using r as an umbral symbol we obtain

$$x^{(r)'}\frac{\partial F}{\partial x^{(r)}} - \frac{d}{dv}\left(x^{(r)'}\frac{\partial F}{\partial x^{(r)'}}\right) + \frac{\partial F}{\partial x^{(r)'}}x^{(r)''} = 0$$

or

$$\frac{dF}{dv} - 2\frac{dF}{dv} = 0$$

showing that F is constant along the geodesics. The constant is now zero instead of unity as it was in the case of the non-minimal geodesics.

Before proceeding to calculate the deflection of the light rays it will be well to prove an often quoted property of them. In a statical gravitational field the time coordinate $x^{(4)}$ does not enter into $F \equiv (ds)^2$. Hence

$$\frac{d}{dv}\left(\frac{\partial F}{\partial x^{(4)'}}\right) = 0 \qquad \text{or} \qquad \frac{\partial F}{\partial x^{(4)'}} = \text{const.}$$

If, now, in the discussion of Ch. 6, § 1, instead of keeping both ends of the " varied curve " $C(\alpha)$ fixed, we had allowed the ends to vary also, the part of δI which came outside the sign of integration when we integrated by parts would not vanish automatically. Since the first variation is to vanish when the end points are fixed as well as when they vary the part under the sign of integration vanishes as before yielding the Eulerian equations but *in*

addition we have the *end condition*

$$\left(\frac{\partial F}{\partial x^{(r)\prime}}\,\delta x^{(r)}\Big|_1^2\right)_{a=0} = 0 \qquad (r \text{ umbral})$$

If now all the coordinates but $x^{(4)}$ are kept fixed

$$\delta x^{(1)} = 0 = \delta x^{(2)} = \delta x^{(3)}$$

and we find since $\dfrac{\partial F}{\partial x^{(4)\prime}}$ is constant over the extremal curve

$$\frac{\partial F}{\partial x^{(4)\prime}}\,\delta x^{(4)}\Big|_1^2 = 0 \qquad \text{or} \qquad \frac{\partial F}{\partial x^{(4)\prime}}\,\delta \int_1^2 dx^{(4)} = 0$$

and as

$$\frac{\partial F}{\partial x^{(4)\prime}} \neq 0$$

we have

$$\delta \int_1^2 dx^{(4)} = 0$$

which is known as the *Fermat* or *Huyghens' Principle of Least Time.* It is an immediate consequence of the absence of $x^{(4)}$ from $(ds)^2$; there is a similar theorem for the symmetrical attracting center:

$$\delta \int_1^2 dx^{(3)} = 0$$

but this has no special utility. The Fermat Principle states that, given two fixed points in space (by fixed is meant that the three space coordinates for an observer attached to the gravitating center are constant), a light signal passes from one to the other in such a way that the first variation of the time interval is a minimum.

With the same notation as that employed in § 3 we find

$$g_3\varphi' = -\,h; \qquad g_4 t' = C$$

where h and C are constants and we find exactly as before that a proper choice of our initial conditions for θ enables us to

write $\theta = \pi/2$. The only difference is that (D') is replaced by

$$(1 - 2m/r)^{-1}(r')^2 + r^2(\varphi')^2 - (1 - 2m/r)(t')^2 = 0$$

whence on writing $r = 1/u$ and using $r^2\varphi' = h$ we find

$$\left(\frac{du}{d\varphi}\right)^2 + u^2 = 2mu^3 + \frac{C^2}{h^2}$$

In order to get an idea of the order of magnitude of the constants C, h of integration we make a first trial-approximation. The largest value that u can have is $1/R$ where R, the radius of the sun, $= 697{,}000$, the units being kilometers. Hence we neglect, for the moment, the u^3 term in comparison with the others and find at once

$$u = \frac{C}{h}\sin(\varphi - \varphi_0)$$

where φ_0 is a constant of integration. Hence C/h is the largest value of u and is therefore a small quantity of the order $1/10^5$. Denoting this small quantity by α (α is the positive square root of C^2/h^2) we have

$$\left(\frac{du}{d\varphi}\right)^2 = 2mu^3 - u^2 + \alpha^2.$$

The discriminant (cf. Ch. 7, § 3) of the cubic on the right is $4\alpha^2(1 - 27m^2\alpha^2)$, a positive quantity, so that the three roots are real. When $\alpha = 0$ they reduce to $1/2m$, 0, 0, so that trying in turn $k\alpha$ and $(1/2m) + k\alpha$ we find the first approximation to the three roots $u_3 = -\alpha$, $u_2 = \alpha$, $u_1 = 1/2m$ where we have arranged the roots so that $u_3 < u_2 < u_1$. For a second approximation, we try in turn $-\alpha + k\alpha^2$, $\alpha + k\alpha^2$, $(1/2m) + k\alpha^2$ for u_3, u_2, u_1 respectively and find

$$u_3 = -\alpha + m\alpha^2; \qquad u_2 = \alpha + m\alpha^2; \qquad u_1 = (1/2m) - 2m\alpha^2.$$

We now, as before, determine a linear transformation which sends $u = u_3$ into $z = 0$; $u = u_2$ into $z = 1$. It is $u = a + bz$

ıere $a = u_3$, $b = u_2 - u_3$ and then the third root $u = u_1$ es into $z = 1/k^2$ where

$$k^2 = \frac{u_2 - u_3}{u_1 - u_3}$$

ow the cubic $2mu^3 - u^2 + \alpha^2$ cannot be negative in our problem ır can u itself. At remote distances from the sun $u \to 0$ so at initially $u = 0$ and it increases to $u = u_2$ at which point u ıs a maximum value, since $\left(\frac{du}{d\varphi}\right) = 0$ there. Then u_2 begins to crease and the radical $\sqrt{2mu^3 - u^2 + \alpha^2}$ in the expression for

$$d\varphi = \frac{du}{\sqrt{2mu^3 - u^2 + \alpha^2}}$$

so changes sign so that $d\varphi$ keeps its sign. The angle ϕ between point at a remote distance and the perihelion of the light ray given by the integral $\int_0^{u_2} \frac{du}{\sqrt{2mu^3 - u^2 + \alpha^2}}$. The excess of vice this over π is the deflection D experienced by the ray. ence

$$D + \pi = 2\int_0^{u_2} \frac{du}{\sqrt{2mu^3 - u^2 + \alpha^2}}$$

hich on writing $u = a + bz$ becomes

$$= \frac{2}{\sqrt{2mb}} \int_{-(a/b)}^{1} \frac{dz}{\sqrt{z(1 - z)\left(\frac{1}{k^2} - z\right)}}$$

n making the final substitution $z = \sin^2\theta$ this becomes

$$D + \pi = \frac{4k}{\sqrt{2mb}} \int_{\sin^{-1}\sqrt{\frac{u_3}{u_3 - u_2}}}^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2\sin^2\theta}}$$

ow

$$k^2 = \frac{2\alpha}{\frac{1}{2m} + \alpha - 3m\alpha^2} = 4m\alpha + \text{higher powers in } \alpha$$

so that, m being 1.47, k^2 is a small quantity of the same order as α. Hence $(1 - k^2 \sin^2 \theta)^{-1/2}$ can be expanded in a rapidly convergent series and an integration of the initial terms of this series gives a high approximation to $D + \pi$. On substituting the values of k and b the multiplier of the integral becomes

$$4\sqrt{\frac{1}{2m(u_3 - u_1)}} = 4(1 + 2m\alpha)^{-1/2} = 4(1 - m\alpha)$$

whilst the lower limit of the integral is

$$\sin^{-1}\sqrt{\frac{1 - m\alpha}{2}} = \sin^{-1}\frac{1}{\sqrt{2}}(1 - \tfrac{1}{2}m\alpha)$$

Here it is necessary to use the second approximation since u_3 is to be *divided* by $u_3 - u_2$ itself a small quantity of the first order. On expanding

$$\sin\left(\frac{\pi}{4} + \epsilon\right) = \frac{1}{\sqrt{2}}\left(1 + \epsilon - \frac{\epsilon^2}{2}\cdots\right)$$

by Taylor's theorem we have for the lower limit $(\pi/4) - \frac{1}{2}m\alpha$ so that

$$D + \pi = 4(1 - m\alpha)\left[\theta + \frac{k^2}{4}(\theta - \sin\theta\cos\theta) + \cdots\right]_{\frac{\pi}{4} - \frac{m\alpha}{2}}^{\pi/2}$$

In the term multiplied by

$$\frac{k^2}{4} = m\alpha + \cdots$$

it is sufficient to take the rough approximation $\pi/4$ to the lower limit and we have

$$D + \pi = 4(1 - m\alpha)\left[\frac{\pi}{4} + \frac{m\alpha}{2} + m\alpha\left(\frac{\pi}{4} + \frac{1}{2}\right)\right]$$
$$= (1 - m\alpha)[\pi + m\alpha(4 + \pi)]$$

giving

$$D = 4m\alpha$$

α, in this expression, is the maximum value u_2 of u (to a first

approximation), i.e., is the reciprocal of the radius of the sun. An idea as to the closeness of this approximation is obtained by using the second approximation

$$\frac{1}{R} \equiv u_2 = \alpha + m\alpha^2$$

The positive root α of this quadratic is

$$\alpha = \left\{ -1 + \left(1 + \frac{4m}{R}\right)^{1/2} \right\} \Big/ 2m = \frac{1}{R}\left(1 - \frac{m}{R}\right)$$

so that writing $\alpha = 1/R$ is equivalent to neglecting m/R in comparison with unity or to a neglect of 1 part in 5.10^5. On substituting $m = 1.47$, $R = 697{,}000$ in the expression $D = 4m/R$ and converting this radian measure into seconds of arc we find the value 1.73″ predicted by Einstein for a light ray which just grazes the sun.*

*For a fuller discussion of the problems dealt with in this chapter reference is made to two papers by the author in the Phil. Mag. of dates Jan. (1922) and March (1922) respectively. For an alternative treatment of the subject matter of §2 the reader should consult the paper *Concomitants of Quadratic Differential Forms* by A. R. Forsyth in the Proc. Roy. Soc. Edin. May (1922).

Date Due

[illegible] 2 4 '52			
APR 1 5 '53			
OCT 2 3 '53			
OCT 3 0 '5[illegible]			
DEC 2 '5[illegible]			
AUG 2 8 '61			
NOV 2 '61			
Nov. 17 '61			
NOV 8 '62			
NOV 21 '7[illegible]			
Demco 293-5			

513.7 M97v

Murnaghan, F. D.

Vector analysis and

the t eory of relativ-

513.7 M97v

The Hunt Library
Carnegie Institute of Technology
Pittsburgh, Pennsylvania